

Northumbria Research Link

Citation: Barrett, Lee (2012) Investigation of the molecular basis of virulence of *Streptococcus equi*. Masters thesis, Northumbria University.

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/8509/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

**Investigation of the molecular basis of
virulence of *Streptococcus equi***

Lee Thomas Barrett

M.Phil

October 2011

**Investigation of the molecular basis of
virulence of *Streptococcus equi***

Lee Thomas Barrett

A thesis submitted in partial fulfilment of the
requirements of the University of Northumbria
at Newcastle for the degree of Master of
Philosophy

Research undertaken in the School of Life
Sciences and in collaboration with The
Animal Health Trust, Newmarket.

October 2011

Abstract

Equine strangles is a purulent lymphadenitis of the head and neck and is caused by *Streptococcus equi* ssp *equi* remaining a worldwide, endemic infection that represents around 30 % of all annually recorded incidents of equine disease. Despite much effort, current vaccination strategies have proved largely ineffective towards *S. equi*, ssp *equi*, with the current focus based on ‘reverse vaccinology’, using genome sequence data of *S. equi* ssp *equi* to identify surface exposed and secreted proteins.

Streptococcus equi shares much of its genome with the genetically closely related, *S. equi* ssp *zooepidemicus*. Despite this close genetic association, *S. equi* ssp *zooepidemicus* exhibits very distinct pathogenicity variations. *S. equi* ssp *zooepidemicus* can infect a wide variety of vertebrate hosts showing a high degree of antigenic variability and a large amount of strain variability whereas *S. equi* is largely equine specific and processes a limited number of strain types. Understanding the molecular basis of virulence of these contrasting organisms therefore remains a key requirement if a suitable vaccination.

Identification of potential virulence factors has been greatly aided by the availability of the full genome sequence of *S. equi* ssp *equi* (4047) and *S. equi* ssp *zooepidemicus* (H70).

In this study, we identified and investigated a range of putative virulence factors including the covalently attached surface protein SEQ2190, the secreted phospholipase enzymes SlaA (SEQ0849) and SlaB (SEQ2155) and finally the lipoprotein, acid phosphatase SeLppC (SEQ0346) and its orthologue SzLppC (SZO16870). In all cases, using molecular biology techniques the coding region representative for each of these proteins was cloning into expression constructs, expressed and purified and further investigations carried out.

Although efforts to obtain a 3-dimensional structure of SEQ2190 were unsuccessful, bioinformatic investigations have identified SEQ2190 as a unique protein in *S. equi* ssp *equi* with a putative structure suggesting a role in bacteria-host interaction.

SlaA was demonstrated to be an active sPLA₂ enzyme, active against a 1, 2-dithio analog of diheptanoyl phosphatidylcholine but not 2-deoxy-2-thio- Arachidonoyl phosphatidylcholine. It was also demonstrated to have a specific requirement for the divalent ion, Ca²⁺ for activity, an optimum temperature higher than expected (40°C), a K_m of 14.40 ± 7.866 mM and a specific activity of 5.06x10⁻² ± 3.01x10⁻³ μmol/min/mg. SlaA was also shown to react to equine post-infection convalescent serum.

Although some characterisation had been previously undertaken regarding whole cell SeLppC extracts, in this study we produced expression constructs of SeLppC and its orthologous pseudogene (SzLppC) found in the genome strain of *S. equi* ssp *zooepidemicus*. We used purified SeLppC and SzLppC to demonstrate activity against pNPP and two biologically significant substrates (5'AMP and 5'UMP) and recorded the K_m values for each. We also identified SeLppC to not possess a specific activity for the divalent ion Cu²⁺, a lower than expected pH optimum and a higher than expected temperature range. Furthermore we also demonstrated that SzLppC is not a pseudogene, indicating an error within the *S. equi* ssp *zooepidemicus* genome strain.

Table of Contents

1. Introduction

1.1.	The Streptococcus	1
1.2.	<i>Streptococcus equi</i> subspecies <i>equi</i> and related taxa	3
1.3.	Equine Strangles	5
1.4.	The Gram-positive cell envelope	11
1.4.1.	Transport through the cytoplasmic membrane	15
1.4.2.	Covalent attachment of proteins to the cell wall	17
1.4.3.	Non-covalent attachment of proteins to the Gram-positive cell wall	20
1.4.4.	Lipoprotein biosynthesis	21
1.5.	The molecular basis of virulence	23
1.6.	Virulence factors of <i>Streptococcus equi</i> sub-species and related pathogens	24
1.6.1.	Fibrinogen-binding proteins	24
1.6.1.1.	M-protein	24
1.6.1.2.	H-binding protein	27
1.6.2.	Hyaluronic acid capsule	28
1.6.3.	Lipoproteins	30
1.6.4.	Superoxide dismutase	32
1.6.5.	Super antigens	32
1.6.6.	Fibronectin-binding proteins	34
1.6.7.	Collagen-binding proteins	36

1.6.8.	Immunoglobulin G (IgG) endopeptidases	37
1.7.	Vaccines against <i>S. equi</i>	38
1.8.	Phospholipases	41
1.8.1.	Phospholipase putative roles in bacterial virulence	43
1.8.2.	The mammalian arachidonic acid cascade	46
1.8.3.	PLA ₂	48
1.8.4.	sPLA ₂ of Streptococci species	49
1.9.	Acid Phosphatase	52
1.9.1.	Bacterial non-specific acid phosphatase (NSAP)	53
1.9.1.1.	Class A and B NSAP	53
1.9.1.2.	Class C NSAP	54
1.9.2.	Putative NSAP role in virulence	55
2.	Materials and methods	
2.1.	Chemicals	57
2.2.	Commercial Kits	57
2.3.	Equipment	57
2.4.	Media and Solutions	58
2.5.	Bacterial strains, storage and growth conditions	58
2.6.	<i>E. coli</i> bacterial strains and vectors	58
2.7.	Antibiotics	61
2.8.	Enzymes	62
2.9.	Oligonucleotides	63

2.10.	Spectrophotometry	67
2.11.	Preparation of chemically competent <i>E. coli</i> TOP10	68
2.12.	DNA methods	
2.12.1.	Streptococcal genomic DNA extraction	68
2.12.2.	Restriction enzyme digestion	69
2.12.3.	Polymerase chain reaction	69
2.12.4.	Agarose gel electrophoresis of DNA samples	71
2.12.5.	Cloning of fragments into pCR Blunt	71
2.12.6.	Plasmid isolation	73
2.12.7.	Cloning of fragments into pET28a	74
2.12.8.	DNA sequencing	76
2.13.	General protein methods	
2.13.1.	SDS-PAGE	76
2.13.2.	Protein extraction	79
2.13.3.	Concentration and buffer exchange of proteins	79
2.13.4.	Protein purification by IMAC chromatography	79
2.13.5.	Protein purification by FPLC chromatography	80
2.13.6.	Western Blotting	81
2.13.7.	Protein crystallisation	84
2.14.	Enzyme assay methods	
2.14.1.	General kinetic analysis considerations	84
2.14.2.	sPLA ₂ assay methods	

2.14.2.1.	SIGMA assay	85
2.14.2.2.	Price colourmetric assay	86
2.14.2.3.	Chromogenic substrate release assay	86
2.14.3.	Acid phosphatase assay	89
2.14.4.	Phosphate assay	91
2.15.	Calculations of Km and specific activity	
2.15.1.	sPLA ₂	92
2.15.2.	Acid phosphatase	93
2.16.	Zymography	93
2.17.	Bioinformatic analyses	94
3.	Investigation into the covalently attached protein SEQ2190 of <i>S. equi</i>.	
3.1.	Introduction	96
3.2.	Bioinformatic analysis of SEQ2190	98
3.3.	PCR of <i>seq2180</i> and <i>seq2190</i>	111
3.4.	Cloning of <i>seq2180</i> and <i>seq2190</i> (full length and truncated forms)	119
3.5.	Protein expression	123
3.6.	Protein purification	128
3.7.	Preperation of crystals for X-ray crystallography	135
3.8.	Discussion and future work	137
4.	Investigation into the secreted protein SlaA of <i>S. equi</i>	
4.1.	Introduction	141
4.2.	Bioinformatical analysis of SlaA and an a SlaA homologue in <i>S. equi</i>	143

4.3.	PCR of <i>seq0849</i> and <i>seq2155</i>	149
4.4.	Cloning of <i>seq0849</i> and <i>seq2155</i>	152
4.5.	Protein expression of SEQ0849 and SEQ2155	154
4.6.	Protein purification of SlaA and SlaB	160
4.7.	Confirmation of His ₆ -tagged SlaA and SlaB by western blotting	166
4.8.	Detection of antibodies to SlaA and SlaB in equine convalescent Serum	167
4.9.	Enzymatic assays of SlaA and SlaB activity	
4.9.1.	SIGMA™ method	169
4.9.2.	Price colourmetric method	170
4.9.3.	Chromogenic substrate release assay	174
4.9.3.1.	Michaelis-Menten parameters for SlaA and SlaB	176
4.9.3.2.	Determination of temperature optimum	179
4.9.3.3.	Determination of divalent ion requirement	179
4.9.3.4.	Determination of SlaA and SlaB against arachidonyl thio-PC	184
4.10.	Discussion and future work	186
5.	Investigation into the Acid phosphatase, LppC of <i>S. equi</i> and <i>S. zooepidemicus</i>	
5.1.	Introduction	194
5.2.	Bioinformatic analysis of SeLppC and SzLppC	196
5.3.	PCR across <i>SzLppC</i> Signal Peptide region	202
5.4.	PCR of mature <i>SeLppC</i> and <i>SzLppC</i>	204
5.5.	DNA sequencing of <i>SzLppC</i>	206

5.6.	Identification of LppC enzymes in <i>S. equi</i> and <i>S. zooepidemicus</i>	209
5.7.	Molecular cloning of <i>SeLppC</i> and <i>SzLppC</i>	211
5.8.	Protein expression of SeLppC and SzLppC	213
5.9.	Protein purification of SeLppC and SzLppC	216
5.10.	Confirmation of His ₆ -tagged SeLppC and SzLppC	218
5.11.	Detection of SeLppC and SzLppC using anti-LppC antiserum	218
5.12.	Western immunoblotting of purified SeLppC and SzLppC using equine convalescent serum	219
5.13.	Zymographic detection of acid phosphatase activity	220
5.14.	Enzymatic assays of SeLppC and SzLppC activity	
5.14.1.	Michelis-Menton parameters	225
5.14.2.	Determination of temperature optimum of SeLppC and SzLppC	236
5.14.3.	Determination of divalent ion requirement	236
5.14.4.	Determination of optimum pH	237
5.15.	Discussion and future work	242
6.	Conclusion	249
7.	References	252
i.	Appendices	
a.	Appendix I – Media, solutions and enzyme buffers	289
b.	Appendix II – Commercial kits and equipment	297
c.	Appendix III– Crystal tray screens	301
d.	Appendix IV– Sequencing outputs	313

e. Appendix V– Cloning vector maps, molecular weight markers

and standards

321

Table of Figures

Figure 1.1	Current system of differentiation of pathogenically important streptococci	2
Figure 1.2	Pony with severe swelling of the lymph nodes	8
Figure 1.3	Pony with submandibular lymph node abscess	9
Figure 1.4	The characteristic signs of purpura haemorrhagica	10
Figure 1.5	The peptidoglycan structures from <i>S. aureus</i> and <i>S. pyogenes</i>	13
Figure 1.6	The major surface protein types in Gram-positive bacteria.	14
Figure 1.7	The adhesion of sortase dependent proteins to the Gram-positive bacterial cell wall	19
Figure 1.8	Current classifications of phospholipases	42
Figure 1.9	The AA cascade, indicating the three key enzymatic pathways of AA	47
Figure 1.10	The demonstrated and putative effects of SlaA from <i>S. pyogenes</i> on the eukaryotic cell	51
Figure 2.1	Schematic representation of sPLA ₂ activity reaction	88
Figure 2.2	Schematic representation of acid phosphatase activity reaction	90
Figure 3.1	Alignment of SEQ2190 and SEQ2180	100
Figure 3.2	Alignment of SEQ2180 and the orthologue SZO18890	102
Figure 3.3	Multiple alignments of SEQ2190, SEQ2180 and the orthologue	

	SZO18970	106
Figure 3.4	Predicted secondary structure of mature SEQ2190	109
Figure 3.5	Artemis screen capture of the <i>seq2180</i> and <i>seq2190</i> locus and their respective <i>S. zooepidemicus</i> orthologues <i>szo18890</i> and <i>szo18970</i>	110
Figure 3.6	Genomic DNA extraction of <i>S. equi</i> 4047 and <i>S. zooepidemicus</i> H70	113
Figure 3.7	PCR amplification of <i>seq2190</i> and 3' extension	114
Figure 3.8	PCR amplification of <i>seq2180</i> and 3' extension.	115
Figure 3.9	PCR amplification of <i>seq2190</i>	116
Figure 3.10	PCR amplification of <i>seq2180</i>	117
Figure 3.11	PCR amplification of truncated <i>seq2190</i> encoding SEQ2190 N	118
Figure 3.12	Restriction endonuclease digestion of p2190 plasmid	120
Figure 3.13	Restriction endonuclease digestion of p2180 plasmid	121
Figure 3.14	Restriction endonuclease digestion of p2190 N plasmid	122
Figure 3.15	15% SDS PAGE gel demonstrating SEQ2190 expression trials	124
Figure 3.16	15% SDS PAGE gel demonstrating SEQ2190 expression	125
Figure 3.17	15% SDS PAGE gel demonstrating SEQ2180 expression	126
Figure 3.18	15% SDS PAGE demonstrating SEQ2190 N expression	127
Figure 3.19	15% SDS PAGE of IMAC purified fractions of SEQ2190	129
Figure 3.20	15% SDS PAGE of IMAC purified fractions of SEQ2180	130
Figure 3.21	15% SDS PAGE of IMAC purified fractions of SEQ2190 N	131
Figure 3.22	15% SDS PAGE gel of FPLC purified fractions of SEQ2190	132
Figure 3.23	15% SDS PAGE gel showing FPLC purified fractions of SEQ2180	133

Figure 3.24	15% SDS PAGE gel showing FPLC purified fractions of SEQ2190 N	134
Figure 3.25	Crystal of SEQ2190	136
Figure 3.26	Crystal of SEQ2190	136
Figure 3.27	Proposed predicted secondary structure model of the mature SEQ2190 protein with three distinct structural domains.	140
Figure 4.1	Multiple alignment of SEQ0849, SEQ2155 and the <i>S. zooepidemicus</i> orthologue (SZO18670)	146
Figure 4.2	Pairwise alignment of SlaA and the SpSlaA protein encoded by temperate phage of <i>S. pyogenes</i>	147
Figure 4.3	Identification of Pfam family	148
Figure 4.4	PCR amplification of <i>seq0849</i>	150
Figure 4.5	PCR amplification of <i>seq2155</i>	151
Figure 4.6	Restriction endonuclease digestion of p0849 and p2155	153
Figure 4.7	17% SDS PAGE of SlaA expression trials	156
Figure 4.8	17% SDS PAGE of SlaA expression trials	157
Figure 4.9	17% SDS PAGE of SlaA expression	158
Figure 4.10	17% SDS PAGE of SlaB expression trials	159
Figure 4.11	17% SDS PAGE of IMAC purified fractions of SlaA	161
Figure 4.12	17% SDS PAGE of IMAC purified fractions of SlaB	162
Figure 4.13	17% SDS PAGE of FPLC purified fractions of SlaA	163
Figure 4.14	17% SDS PAGE of FPLC purified fractions of SlaB	164
Figure 4.15	17% SDS PAGE of the purified SlaA and SlaB products obtained	165
Figure 4.16	His ₆ -tag Western blot analyses of SlaA and SlaB	167

Figure 4.17	Western blot analyses of purified SlaA and SlaB against equine convalescent serum (pre and post infection).	168
Figure 4.18	Enzyme assay by the method of Price (2007)	171
Figure 4.19	Enzyme assay by the method of Price (2007)	172
Figure 4.20	sPLA ₂ enzyme assay by the method of Price (2007)	173
Figure 4.21	Standard rate of sPLA ₂ activity over against time	175
Figure 4.22	Rate of SlaA activity against substrate concentration	177
Figure 4.23	Lineweaver-Burk plot of SlaA against DHT-PC	178
Figure 4.24	The effect of temperature on the rate of SlaA enzyme activity against DHT-PC	181
Figure 4.25	The effect of divalent cation replacement on the rate of SlaA enzyme activity	182
Figure 4.26	The effect of calcium ion concentration on the rate of SlaA enzyme activity	183
Figure 4.27	sPLA ₂ specific enzyme activity of positive control phospholipase against different lipid substrates	185
Figure 4.28	Amino acid sequence of SlaA highlighting the effect of amino acid substitution as demonstrated by Nagiec and colleagues (Nagiec et al., 2004)	193
Figure 5.1	Figure 5.1: Pairwise alignment of <i>SeLppC</i> and <i>SzLppc</i>	198
Figure 5.2	Translated amino acid sequence of SzLppC, either side of the stop codon introduced by the putative frameshift	199
Figure 5.3	Pairwise alignment mature of <i>SeLppC</i> and <i>SzLppc</i>	200

Figure 5.4	PCR amplification of <i>SzLppC</i> signal peptide region	203
Figure 5.5	Multiple alignment of sequenced portion of <i>SzLppC</i> (this study, 4 replicates) and the deposited nucleotide sequence of <i>SzLppC</i> , obtained from the <i>S. zooepidemicus</i> H70 genome	205
Figure 5.6	15% SDS PAGE of range of <i>S. equi</i> and <i>S. zooepidemicus</i> strains	207
Figure 5.7	LppC Western blot analysis of a range of <i>S. equi</i> and <i>S. zooepidemicus</i> Strains	208
Figure 5.8	PCR amplification of <i>seq0346</i> and <i>szo16870</i>	210
Figure 5.9	Restriction endonuclease digestion of pSeLppC and pSzLppC	212
Figure 5.10	15% SDS PAGE gel demonstrating SeLppC and SzLppC expression trials	214
Figure 5.11	15% SDS PAGE demonstrating SeLppC and SzLppC expression	215
Figure 5.12	15% SDS PAGE demonstrating proteins used in further work in this study	217
Figure 5.13	His ₆ -tag western blot analyses of SeLppC and SzLppC	220
Figure 5.14	Anti-LppC Western blot analysis of purified SeLppC and SzLppC products	221
Figure 5.15	Western blot analysis of purified SeLppC and SzLppC against equine convalescent serum (pre and post infection)	222

Figure 5.16	15% SDS PAGE Zymograph of SeLppC and SzLppC and SDS PAGE (Standard Coomassie stained replicate)	224
Figure 5.17	Standard absorbance of control potato acid phosphatase against NPP	226
Figure 5.18	Lineweaver-Burk plot of control potato acid phosphatase against <i>p</i> NPP	227
Figure 5.19	Lineweaver-Burk plot for SeLppC against <i>p</i> NPP	228
Figure 5.20	Lineweaver Burk plot for SzLppC against <i>p</i> NPP	229
Figure 5.21	Lineweaver Burk plot for SeLppC against 5'UMP	230
Figure 5.22	Lineweaver-Burk plot for SzLppC against 5'UMP	231
Figure 5.23	Lineweaver Burk plot of SeLppC against 5'AMP	232
Figure 5.24	Lineweaver Burk plot of SzLppC against 5'AMP	233
Figure 5.25	Standard curve of inorganic phosphate assay	234
Figure 5.26	The effect of temperature on the rate of SeLppC and SzLppC enzyme activity.	238
Figure 5.27	The effect of divalent ions on the rate of SeLppC enzyme activity	239
Figure 5.28	The effect of divalent ions on the rate of SzLppC enzyme activity	240
Figure 5.29	The effect of changes in pH on the rate of SeLppC and SzLppC enzyme activity	241

List of Tables

Table 2.1	Bacterial strains and plasmids used in this study	59
Table 2.2	Antibiotics used in this study	61
Table 2.3	Enzymes and co-constituents	62
Table 2.4	Nucleotide primers used in this study	64
Table 2.5	Recombinant plasmids generated in this study	66
Table 2.6	PCR reaction conditions	70
Table 2.7	Components of PCR-Blunt ligation reaction	72
Table 2.8	Components of pET vector ligation	75
Table 2.9	Composition of SDS-PAGE stacking gel	77
Table 2.10	Composition of SDS-PAGE resolving gel	78
Table 2.11	Primary antibodies used in this study	82
Table 2.12	Secondary antibodies used in this study	83
Table 3.1	Significant orthologues of the mature SEQ2190 protein	107
Table 4.1	Representative homologues of SEQ0849	145
Table 5.1	Significant orthologues of mature SeLppC	201
Table 5.2	K _m (mM) activities of SeLppC and SzLppC	235
Table 5.3	Recorded K _m 's of SeLppC and SzLppC for three substrates and other closely related NSPC's	248

Acknowledgements

First and foremost, I would like to thank Prof. Iain Sutcliffe and Prof. Gary Black for all of the help, support, guidance and advice they have given me during this project.

I would also like to thank Dr. Andrew Waller for his assistance and the Animal Health Trust for contributing to the funding of this work.

I would also like to show my gratitude to Dr. Alistair Brown assisted in giving me the opportunity to fund myself while writing this thesis and the countless number of colleagues from lab A307/321 both past and present. Especially I would like to thank Paul Purcell who remains the only individual I've ever met to share the same sense of humour as me and Claire Jennings for her valuable support.

I would also like to show my appreciation for the teachings of the Buddha whose wisdom and guidance has provided me with comfort during the many difficult times I experienced during the production of this work.

Finally I wish to thank most of all, my grandfather Thomas (Tot) Gill. Without my grandfathers constant support, advice and humour I would have never made it through these long and difficult years.

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work.

Name:

Signature:

Date:

Abbreviations

5'AMP	Adenoside-5'-monophosphate
5'UMP	Uridine-5'-monophosphate
ABC	ATP-binding cassette
AA	Arachidonic acid
AP	Acid phosphatase
APS	Ammonium persulphate
AT-PC	2-deoxy-2-thio- Arachidonoyl phosphatidylcholine
ATP	Adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
BSA	Bovine serum albumin
CaCl ₂	Calcium chloride
CDS	Coding sequences
CuSO ₄	Copper(II) sulfate
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DHT-PC	1, 2-dithio analog of diheptanoyl phosphatidylcholine
DTNB	5, 5'-dithio- <i>bis</i> -(2-nitrobenzoic acid)

EDTA	2-({2-[bis(carboxymethyl)amino]ethyl} (carboxymethyl)amino)acetic acid
ELISA	Enzyme-linked immunosorbent assay
FPLC	Fast Protein Liquid Chromatography
Fn	Fibronectin
G	Graviton
<i>gluU</i>	N-acetyl-glucosamine-1-phosphate uridyltransferase
HA	Hyaluronic acid
HAP	Hyaluronate-associated protein
<i>hasA</i>	Hyaluronate synthase
<i>hasB</i>	UDP-glucose dehydrogenase
<i>hasC</i>	Glucose-1-phosphate uridyltransferase
HCl	Hydrochloric acid
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
H ₂ O	Water
IL-8	Interleukin 8
IMAC	Protein purification by Immobilized Metal Affinity Chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside

KCl	Potassium chloride
KH ₂ PO ₄	Monopotassium phosphate
kHz	kilohertz
Lgt	Prolipoprotein diacylglyceryl transferase
Lnt	Lipoprotein N-acyl-transferase
Lsp	Lipoprotein signal peptidase
mA	Milliampere
MHC	Major histocompatibility complex
mg	Miligram
MgCl ₂	Magnesium chloride
ml	Mililitre
MLST	Multilocus sequence typing
MSCRAMM's	Microbial surface components recognising adhesive matrix molecules.
Nag	<i>N</i> -acetylglucosamine
NBT-BCIP	Nitro Blue Tetrazolium/5-Bromo-4-Chloro-3-Indolyl Phosphate
Nam	<i>N</i> -acetylmuramic acid
NaOH	Sodium hydroxide

nm	Nanometer
NSAP	Bacterial nonspecific acid phosphatase
OMPLA	Outer membrane phospholipase A
PAF	Platelet activating factor
PBS	phosphate buffered saline
PBST	phosphate buffered saline and Tween 80
PCR	Polymerase Chain Reaction
<i>pgi</i>	Glucose-6-phosphate isomerise
PL	Phospholipase
PLA	Phospholipase A
PLA ₁	Phospholipase A ₁
PLA ₂	Phospholipase A ₂
PLB	Phospholipase B
PLC	Phospholipase C
PLD	Phospholipase D
<i>p</i> NPP	<i>Para</i> -nitrophenylphosphate
rpm	Revolutions per minute
sAg	Super-antigen

SBP	Substrate-binding protein
Sec pathway	Secretory pathway
<i>SeLppC</i>	<i>LppC</i> of <i>Streptococcus equi sub sp equi</i>
<i>S. equi</i>	<i>Streptococcus equi</i> subspecies <i>equi</i>
SFS	Secreted fibronectin binding protein
SOD	superoxide dismutase
SPase	Signal peptidase
SpSlaA	<i>S. pyogenes</i> sPLA ₂
sPLA ₂	Secreted phospholipases A ₂
<i>S. ruminatorum</i>	<i>Streptococcus equi</i> subspecies <i>ruminatorum</i>
<i>SzLppC</i>	<i>LppC</i> of <i>Streptococcus equi sub sp zooemidemicus</i>
<i>S. zooepidemicus</i>	<i>Streptococcus equi</i> subspecies <i>zooepidemicus</i>
TAT pathway	Twin arginine pathway
TEMED	Tetramethylethylenediamine
TNF α	Tumour necrosis factor
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
μg	Microgram
μl	Microlitre

V

Volt

ZnCl_2

Zinc chloride

1. Introduction

1.1. The genus *Streptococcus*

Differentiation of the genus *Streptococcus* was initially made by Schottmuller in 1903 using haemolysis of blood agar. Subsequently, simple fermentation and tolerance tests were later developed to distinguish many of the streptococci (Facklam 2002). Lancefield later significantly improved the means of streptococcal differentiation with the discovery of specific carbohydrate ‘group’ antigens associated with specific strains of *Streptococcus* (Lancefield 1933). Sherman then expanded on the work of Schottmuller in 1937, proposing the use of the reactions some streptococci produced on blood agar plates (Sherman 1937). This led to streptococci species being placed into either an alpha-haemolytic, beta-haemolytic or non-haemolytic group although in practice haemolysis can be variable (Figure 1.1).

The genus *Streptococcus* was eventually split into three genera, *Enterococcus*, *Lactococcus* and *Streptococcus* as defined in *Bergey’s Manual of Determinative Bacteriology* in 1986. Currently, simple differentiation of streptococcal species is based on an amalgamation of haemolytic phenotype on blood agar plates and Lancefield serotyping. For example, *Streptococcus pyogenes* is regarded as a β -haemolytic, Lancefield Group A *Streptococcus* (Figure 1.1).

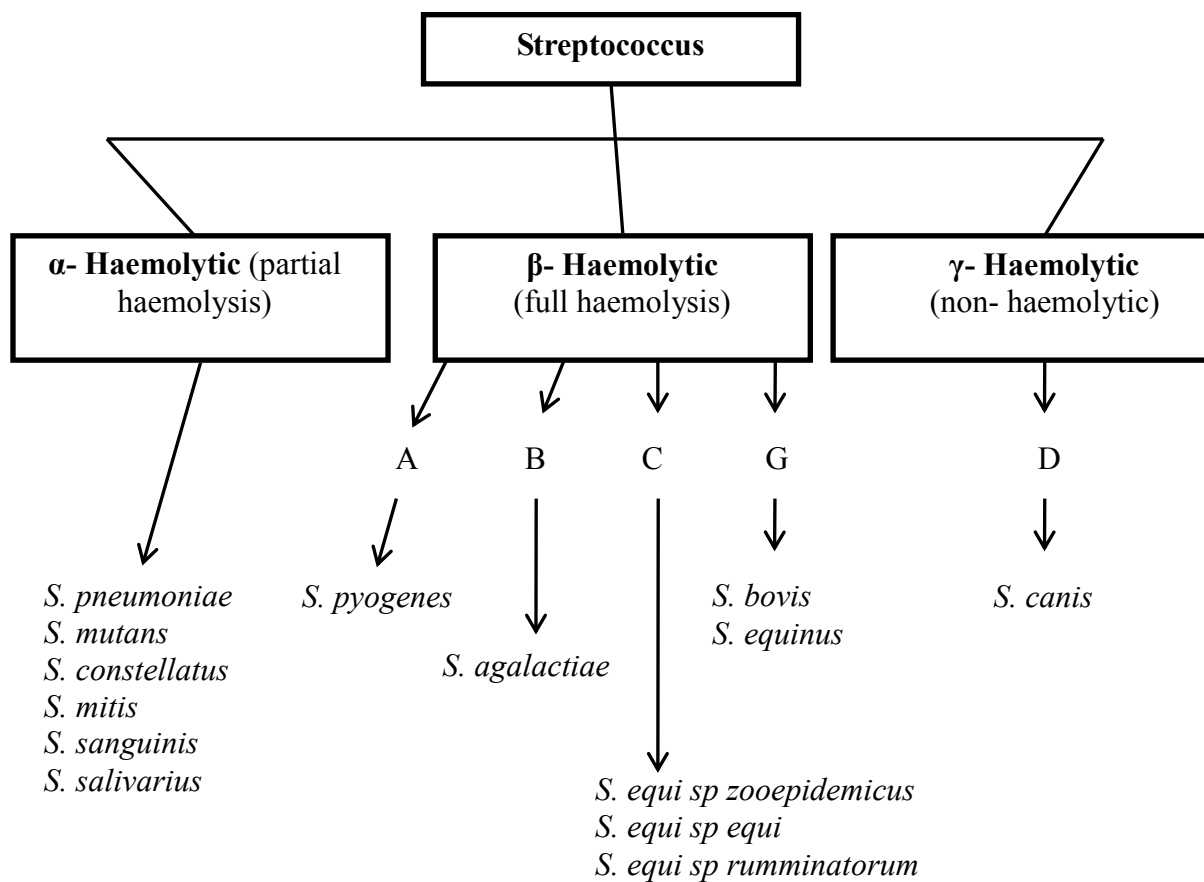


Figure 1.1: Current simplified system of differentiation of pathogenically important streptococci.

Note that many non-haemolytic ‘streptococci’ have since been reclassified as members of the genus *Enterococcus*.

1.2. *Streptococcus equi* subspecies *equi* and related taxa

Streptococcus equi is a β -haemolytic, Lancefield group C *Streptococcus* and consists of three subspecies of zoonotic agents including *Streptococcus equi* subspecies *equi* (herein *S. equi*), *Streptococcus equi* subspecies *zooepidemicus* (herein *S. zooepidemicus*) and *Streptococcus equi* subspecies *ruminatorum* (herein *S. ruminatorum*). *S. equi* is the causative agent of equine strangles, an important worldwide infection in equine species. *S. equi* is regarded as a subtype of *S. zooepidemicus* (Harrington *et al.* 2002) based on multilocus enzyme electropherotypes (Jorm, *et al.* 1994) and also due to comparisons made regarding the 16S-23S rRNA gene intergenic spacers of *S. equi* and *S. zooepidemicus* (Chanter *et al.* 1997). Multilocus sequence typing (MLST) has also revealed this close genetic relationship of *S. equi* and *S. zooepidemicus* (Webb *et al.* 2008).

S. equi shares much of its genome with the genetically closely related *S. zooepidemicus*, with 1671 *S. equi* predicted coding sequences (CDS) having *S. zooepidemicus* orthologues (Holden *et al.* 2009). Both genomes are similar in size with that of *S. equi* (strain 4047) 2,253,793 bp in length compared to *S. zooepidemicus* (strain H70) which is 2,149,866 bp in length. Interestingly, *S. equi* (4047) has a larger number of predicted coding sequences (2,137) than *S. zooepidemicus* (H70). From the remaining 466 non-orthologues *S. equi* CDS the majority (422) are found on mobile genetic elements, accounting for the increase in the *S. equi* genome size (Holden, *et al.* 2009). Additionally *S. equi* displays a larger number of partially deleted genes and pseudogenes than *S. zooepidemicus* indicating a degree of functional loss (Holden, *et al.* 2009). Analysis of the genome strains of *S. equi*, *S. zooepidemicus* and *S. pyogenes* also provides evidence of horizontal genetic exchange occurring between the three species that has influenced the pathogenicity of these bacteria

(Holden *et al.* 2009). The two genome sequenced *S. equi* sub-species also share many structural proteins with *S. pyogenes*, sharing 80% DNA identity (Holden *et al.* 2009).

Despite a large degree of genetic homology the three Group C streptococcal subspecies have many distinct variations regarding each organism's pathogenicity, the molecular basis of which therefore remains an important area of study. Polymorphism of DNA restriction fragments also suggests that there is some limited genomic sequence variability (Takai *et al.* 2000). Populations of *S. zooepidemicus* are extremely diverse with at least 219 sequence types identified. Strains of *S. equi* in comparison are highly homogenous (Kelly *et al.* 2006). MLST of a large collection of isolates identified *S. equi* strains to be either ST-151 or ST-179 (Webb *et al.* 2008).

S. equi is antigenically conserved and is highly host adapted to equids where as the antigenically diverse, *S. zooepidemicus* can be found in many vertebrate hosts including horses (Timoney *et al.* 2007). *S. zooepidemicus* also persists indefinitely as a mucosal and tonsillar commensal. Often a predisposing factor is required prior to *S. zooepidemicus* infection (Timoney *et al.* 2007). Although rare, several cases of human *S. zooepidemicus* infections have been recorded, including a large outbreak of glomerulonephritis in Nove Serrana, Brazil caused by unpasteurised cheese (Baiter *et al.* 2000). Infections of *S. zooepidemicus* in humans have also been implicated in cases of meningitis (CDC 1983) and sepsis (Edwards *et al.* 1988).

S. ruminantium is a relatively newly classified organism and awaiting further study (Fernandez *et al.* 2005) but has also been shown to cause infections in humans (Marchandin *et al.* 2007).

1.3. Equine Strangles

Equine strangles is a purulent lymphadenitis of the head and neck and is caused by *S. equi* (Timoney *et al.* 2007) remaining a worldwide, endemic infection that represents around 30 % of all annually recorded incidents of equine disease (Timoney 2004). Strangles disease is regarded as of significant welfare and economic importance (Guss *et al.* 2009).

Clinical characteristics of equine strangles include fever, cough and pyrexia. This is followed by profuse nasal discharge (Waller and Jolley 2007a) and enlargement/ulceration of the submandibular lymph nodes (Figure 1.2) (Laus *et al.* 2007) as well as abscessation (Figure 1.3). This increased lymph node swelling gives rise to the term ‘strangles’ as in severe cases it can result in restriction of the airways (Waller and Jolley 2007a). Initially *S. equi* enters the host via the nose or mouth, binding to cells on the tonsillar crypts and ventral surface of the soft pallet (Timoney 1999). Infection of the nasopharynx spreads rapidly to the lymph nodes (Timoney 2004) where bacterial multiplication occurs unhindered by polymorphonuclear leukocytes infiltration (Harrington, *et al.* 2002). Lymphadenitis then develops into abscessation.

Bacteria can disseminate to organs distinct from the equine respiratory systems either hematogenously or via lymphatic channels resulting in metastasis, often referred to as ‘bastards strangles’ (Timoney 2004). For example, metastasis of a *S. equi* infection to the brain has been reported (Evers 1968). Purpura haemorrhagica can also be triggered due to infections of *S. equi* (Pusterla *et al.* 2003), which is an immune complex disease and capable of causing oedema (Figure 1.4). The occurrence of oedema in an infected horse is often fatal (Waller and Jolley 2007a).

The recognised site of prolonged bacterial carriage is the guttural pouch (Timoney 2004). This organ has an elusive role but is thought to be implicated in brain-cooling mechanisms in the horse by ventilating and cooling internal carotid arteries (Baptiste 1998; Baptiste *et al.* 2000; Parillo, *et al.* 2009). Prolonged bacterial carriage within the guttural pouch can eventually result in the formation of chondroids, which are discrete bodies formed by the drying and hardening of exudates (Waller and Jolley 2007b). An indicator of subclinical inflammation is often the presence of chondroids within the guttural pouch. The high quantities of *S. equi* bacteria within chondroids suggests that a continual stimulation of the host immune system occurs in the guttural pouch (Verheyen *et al.* 2000).

Common morbidity and mortality rates during an infection outbreak amongst a population are up to 100 % and 10 % respectively (Laus *et al.* 2007).

A variety of mechanisms can allow transmission of *S. equi* amongst susceptible populations, many of which do not require direct horse-horse contact. *S. equi* can persist in standing water for several weeks but has limited survivability in soil and on pasture (Timoney 1999). Incubation can be between four to five days or anything up to 14 days depending on population susceptibility (Timoney 2004). Nasal shedding of bacteria begins after a few days following the onset of pyrexia and can persist for several weeks. Intermittent shedding can occur for several years after infection onset (Timoney 2004). The presence of asymptomatic carriers within susceptible horse populations gives rise to an infection that is endemic in character (Laus *et al.* 2007). Carriers can carry and shed viable infectious quantities of *S. equi* for between seven and thirty nine months (Newton *et al.* 1997). This leads to much difficulty, as it cannot always be determined whether or not an out-break associated horse has the potential to develop the disease or indeed transmit the infection to other horses (Waller and Jolley 2007b).

An efficient method of diagnosis is crucial when dealing with any outbreak of strangles in a susceptible population. Previous diagnostic procedures required utilising traditional microbiological techniques and Polymerase Chain Reaction (PCR) screening of nasopharyngeal swabs of exposed horses (Timoney and Artiushin 1997). A recently developed diagnostic tool is an enzyme-linked immunosorbent assay (ELISA) based detection of serum antibodies that target a full-length protein, SeM (IDEXX™). However, ELISA based testing can be difficult to interpret due to overlapping detection breakpoints in normal and convalescent horses (Davidson *et al.* 2008). Regarding potential carriers, endoscopic examination to confirm chondroids and the sampling of the guttural pouch content remains the only reliable means of identification (Timoney 2004).

S. equi is susceptible to a wide spectrum of current antibiotics in use. However, opinion on whether antibiotic treatment is effective in combating strangles remains a divisive issue (Sweeney *et al.* 2005). The use of antibiotics during the early acute phase of infection may prevent abscess formation. However, the development of protective immunity is often inhibited (Piche 1984), leading to susceptibility to further infections remaining (Sweeney *et al.* 2005). Additionally during treatment, relapses are common. It has been hypothesised that the reason for this is due to ineffective antibiotic penetration to therapeutic levels (Harrington, *et al.* 2002). However, the majority of animals do recover from the disease and are able to eliminate *S. equi* within a period of 4-6 weeks. Complications caused by a strangles infection include metastatic spread of infection, myositis, muscle infarctions and agalactia (Timoney 2004).



Figure 1.2: Pony with severe swelling of the lymph nodes

(Image obtained from The British Horse Society, <http://www.bhswm.org.uk>)



Figure 1.3: Submandibular lymph node abscess

(Image obtained from The Animal Health Trust, http://www.aht.org.uk/bact_intro.html)



Figure 1.4: The characteristic signs of purpura haemorrhagica

Formation of an immune complex resulting in haemorrhage in the mucous membrane.

(Image obtained from The Animal Health Trust, http://www.aht.org.uk/bact_intro.html)

1.4. The Gram-positive cell envelope

Morphologically, Gram-stain positive ('monoderm') bacteria consist of three distinct cellular compartments: the cytosol, a single cytoplasmic membrane, and the surrounding cell envelope (Giesbrecht *et al.* 1976; Sutcliffe 2010). Gram-positive bacteria are differentiated from Gram negative bacteria by having a thicker peptidoglycan wall (also known as murein) and lack a defined outer membrane (Scott and Barnett 2006).

The peptidoglycan is a large macromolecule 15-30 nm thick that encases the cell. The glycan of peptidoglycan is a repeating disaccharide of *N*-acetylmuramic acid linked to *N*-acetylglucosamine by a β -1,4 glycosidic bond (Navarre and Schneewind 1999; Scott and Barnett 2006). The glycan strands vary in length (between 5 to 30 subunits) in relation to the bacterial species and the complete peptidoglycan is formed by the cross-linking of stem peptides and a peptide cross-bridge (Navarre and Schneewind 1999; Snowden and Perkins 1990). In some cases the stem peptides are not linked by peptide-cross bridges but instead via an amide bond as demonstrated by *Listeria monocytogenes*. The exact composition of peptidoglycan is variable based on the organism (Schleifer 1973) (Figure 1.5) .

The primary function of peptidoglycan is in providing a physical protective barrier against osmotic and mechanical stress (Dramsi *et al.* 2008; Navarre and Schneewind 1999). However peptidoglycan has a further primary function in that it provides a scaffold for the anchorage of external structures including secondary cell wall polymers such as teichoic and teichuronic acids and also for the attachment of surface proteins and adhesions (Dramsi *et al.* 2008; Salton *et al.* 1994; Sugiyama *et al.* 2002). The basis of bacterial species and strain specific properties as well as bacteria-host interactions in the case of pathogenic bacteria is

due in many ways to the collective action of these molecular peptidoglycan ‘decorations’ (Marraffini, et al. 2006). Bacteria have evolved several unique mechanisms in which to transport and facilitate the attachment of these proteins, to be discussed in later sections. This includes covalent and non-covalent attachment as well as lipoproteins (figure 1.6).

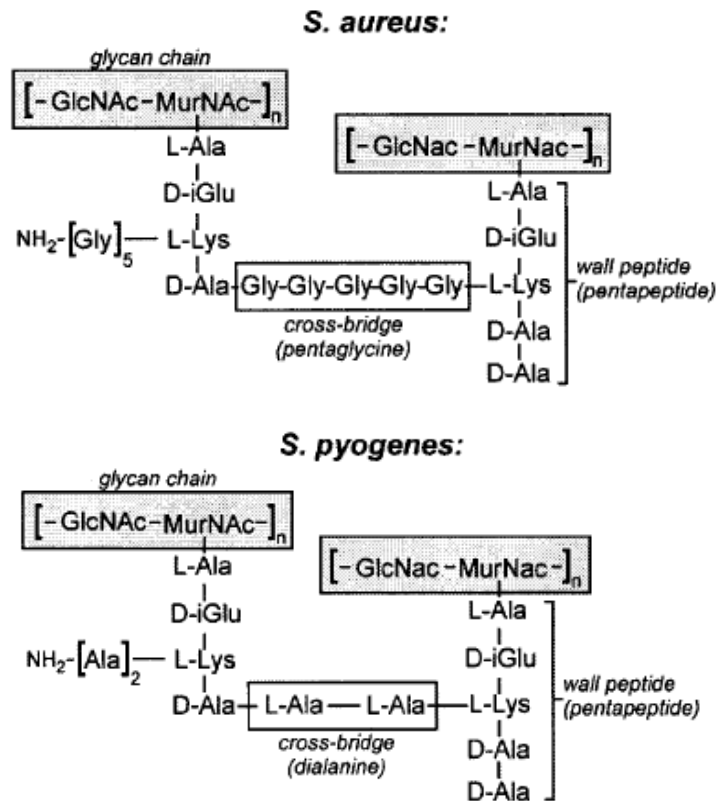


Figure 1.5: The peptidoglycan structures from *Staphylococcus aureus* and *S. pyogenes*.

The glycan chains are composed of a repeating disaccharide, *N*-acetylmuramic acid and *N*-acetylglucosamine. These are linked through the lactyl moiety of *N*-acetylmuramic acid to short wall peptides. The linking of adjacent wall peptides is usually through crossbridge peptides (pentaglycine in *S. aureus* or dialanine in *S. pyogenes*). Figure taken from Navarre and Schneewind (1999).

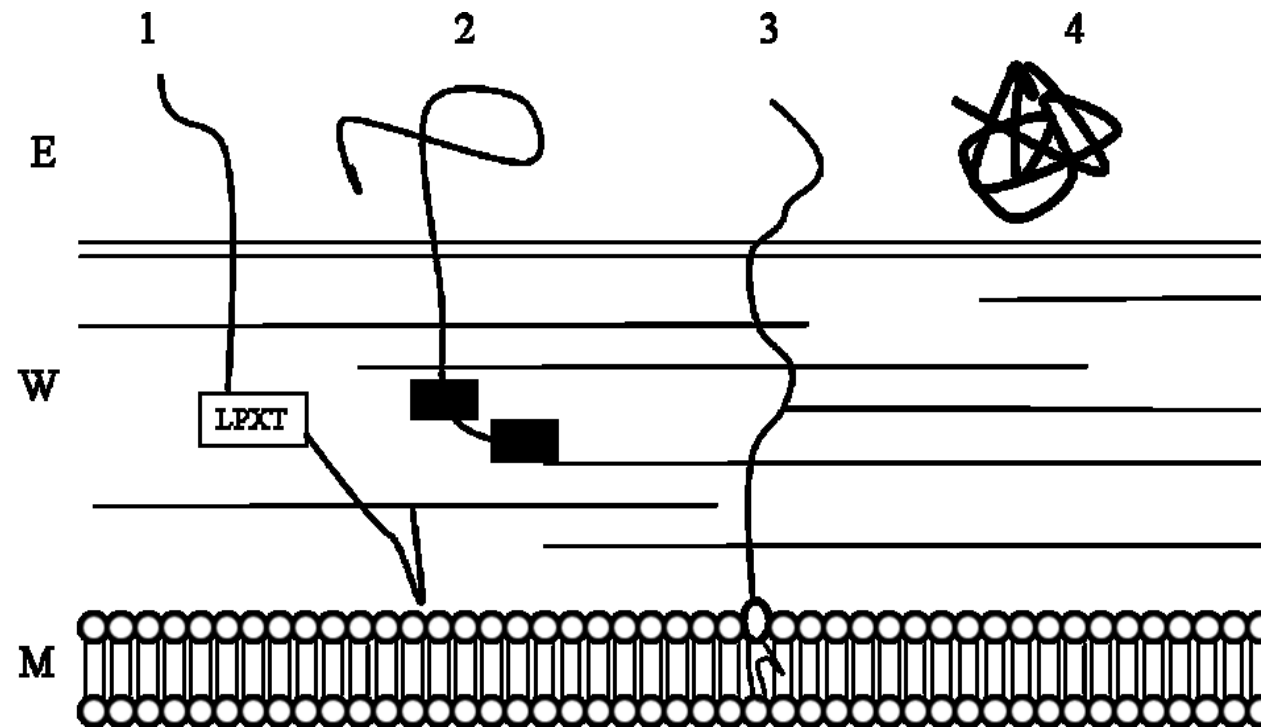


Figure 1.6: The major surface protein types in Gram-positive bacteria.

(1) Covalently bound to the cell wall. (2) noncovalently bound to the cell wall. (3) Membrane anchored lipoprotein. (4) Secreted; M, cell membrane; W, cell wall; E, surface exposed or secreted. Adapted from Barinov et al (2009).

1.4.1. Transport through the cytoplasmic membrane

Bacterial cell integrity is maintained by the separation of the cell cytoplasm from the external environment by a phospholipid membrane. Proteins which function outside the cytoplasm but are synthesized in the cytoplasm therefore need a method of transport through this barrier without compromising cell integrity (de Keyzer *et al.* 2003). The initial step in protein localisation is the transport of the proteins through or into the cell membrane, with most bacterial proteins transported unfolded (Scott and Barnett 2006). The most common mechanism of protein transport is the secretory (Sec) pathway. Further methods of transport include the twin arginine pathway (TAT), although this seems to not be a major pathway in streptococci (Dilks *et al.* 2003) and some specialised proteins also utilise specialist pathways for their transport (Scott and Barnett 2006).

The Sec pathway was originally identified in *Escherichia coli* (Bieker *et al.* 1990; Schatz and Beckwith 1990). The Sec pathway consists of an ATPase, known as SecA (Oliver and Beckwith 1981), and a protein-conducting channel which acts as a large integral membrane domain. The membrane channel is formed by the interaction of two cell membrane proteins, SecY (Emr *et al.* 1981) and SecE (Riggs *et al.* 1988) and their interactions with SecG (Nishiyama *et al.* 1993). SecA enables the use of Adenosine triphosphate (ATP) as an energy source to drive the Sec pathway and when associated with other Sec proteins are referred to as a translocon (Du Plessis *et al.* 2011). Recent studies have also proposed that translocons in Gram-positive bacteria are clustered in restricted regions of the cell membrane known as an ExPortal (Rosch and Caparon 2005) or in other domains (Buist *et al.* 2006). The Sec pathway has two substrate dependent modes of operation: the

insertion of transmembrane segments through a lateral opening and a transversal opening, allowing secretory proteins across the membrane to the periplasm (Facey and Kuhn 2010)

The presence of an N-terminal signal sequence (signal peptide) and Type I signal peptidases (SPase) are a requirement for secreted proteins to use the Sec pathway. Type I Spase is an essential membrane bound endopeptidase, which cleaves the amino-terminal signal peptide from an exported pre-protein (Tuteja 2005). This can be either during or after translocation of the protein has occurred across the membrane (van Roosmalen *et al.* 2004). Following cleavage the mature protein is released from the membrane. Signal peptides consist of three regions: a region at the N-terminus with a net positively charged stretch of polar residues (N-region); a hydrophobic core (H-region) which consists of 10-15 amino acid residues; and finally a region adjacent to the H-region (C-region) which contains the recognition site for Spase I. The H-region is thought to adopt a α -helical conformation (Nielsen *et al.* 1997) where as the C-region is proposed to have a β -stranded conformation (Paetzel *et al.* 1998).

Signal peptides currently have several characterised roles. Signal peptides act as topological determinants for membrane pre-proteins, initiating translocation of pre-protein hydrophilic C-terminal regions. They also ensure that the hydrophobic N-terminal regions of proteins remain in the *cis* orientation at the membrane during translocation (Andersson *et al.* 1992). Signal peptides can also prevent the activation of harmful secretory enzymes in the cell by preventing the folding of nascent chains (Vonheijne and Abrahmsen 1989). Gram-positive bacteria have been shown to have signal peptide sequences that are longer and more hydrophobic than their Gram-negative counterparts (Park *et al.* 1988). Type I Spases are approximately 200 amino acids in size in Gram-positive organisms, compared to around 300 amino acids in Gram-negative organisms. The only known exception to this is the Type I

Spases from *Streptomyces* and *Mycobacterium*, which resemble those of Gram-negative bacteria in size (van Roosmalen *et al.* 2004).

Most bacteria contain at least a single chromosomal copy of the SPase I gene, however many Gram-positive bacteria also encode multiple type I Spases. *Bacillus subtilis* for example has five chromosomally encoded type I SPases (Tjalsma *et al.* 1997). Additional type I Spases can have either a complete or partial redundancy, an example of which can be found in *Staphylococcus aureus* where *spsA* has been found to be redundant (Cregg *et al.* 1996). Regarding bacteria with only one chromosomally encoded type I SPases, this enzyme is essential for the processing of all secretory pre-proteins (van Roosmalen *et al.* 2004). However with those bacteria that possess multiple type I SPases, it has been shown that none of the enzymes by themselves are essential for cell viability (Palacin *et al.* 2002). It is unknown why this is the case but it has been proposed that SPase redundancy may be an advantage for these organisms (van Roosmalen *et al.* 2004).

1.4.2. Covalent attachment of the proteins to the cell wall

Covalent linkage of proteins to the peptidoglycan initially emerged from investigations of protein A of *S. aureus* (Sjoquist *et al.* 1972). Following translocation through the membrane by the Sec pathway, surface proteins are attached covalently to the peptidoglycan wall and are anchored through their carboxyl terminus by membrane-bound thiol transpeptidases called sortases (Mazmanian *et al.* 1999). Surface proteins that are sortase attached also possess a cell wall anchoring domain at their C-termini, constituting a sortase recognition sequence, known as the cell-wall sorting signal (Mandlik *et al.* 2008). This amino acid sequence is often Leu-Pro-X-Thr-Gly (LPXTG) and is followed by a hydrophobic stretch of

amino acids known as a membrane-spanning domain and a positively charged tail (Fischetti *et al.* 1990). Although this motif is common some putative covalently attached proteins of streptococci apparently have displayed an atypical sorting motif (Egan, *et al.* 2010). During the initial process of sortase-mediated attachment, the N-terminal signal sequence is cleaved with the C-terminal hydrophobic domain retaining the protein in the membrane, allowing the LPXTG motif to be recognised. The LPXTG motif is cleaved by sortase between the threonine (T) and glycine (G) and catalyses the formation of an amide bond between the carboxyl group of the threonine and the amino group of peptidoglycan precursor peptides (Jovel *et al.* 2006; Navarre and Schneewind 1994), thus covalently linking the mature protein to the peptidoglycan precursor, which is ultimately incorporated into the growing peptidoglycan (Pallen *et al.* 2001) (Figure 1.7).

Almost all Gram-positive bacteria possess sortase-like proteins, along with two Archaea, *Methanobacterium thermoautotrophicum* (Pallen *et al.* 2001) and *Methanopyrus kandler* (Slesarev *et al.* 2002), which have also been found to encode putative sortase-like proteins. Bioinformatical analysis has also shown that the majority of Gram-positive bacteria also possess additional sortases, and several sortase-like proteins can be found in some individual species (Navarre and Schneewind 1999). These sortases have been shown to be involved in a variety of processes such as pilus assembly, iron acquisition and sporulation (Mandlik *et al.* 2008). A large majority of important surface proteins including virulence-related MSCRAMM (microbial surface components recognising adhesive matrix molecules) and enzymes are covalently anchored in this way to the cell wall in Gram-positive bacteria (Pallen *et al.* 2001). As such, sortase-like proteins have since become an interesting therapeutic agent target in many pathogenic bacteria (Matoba *et al.* 2002).

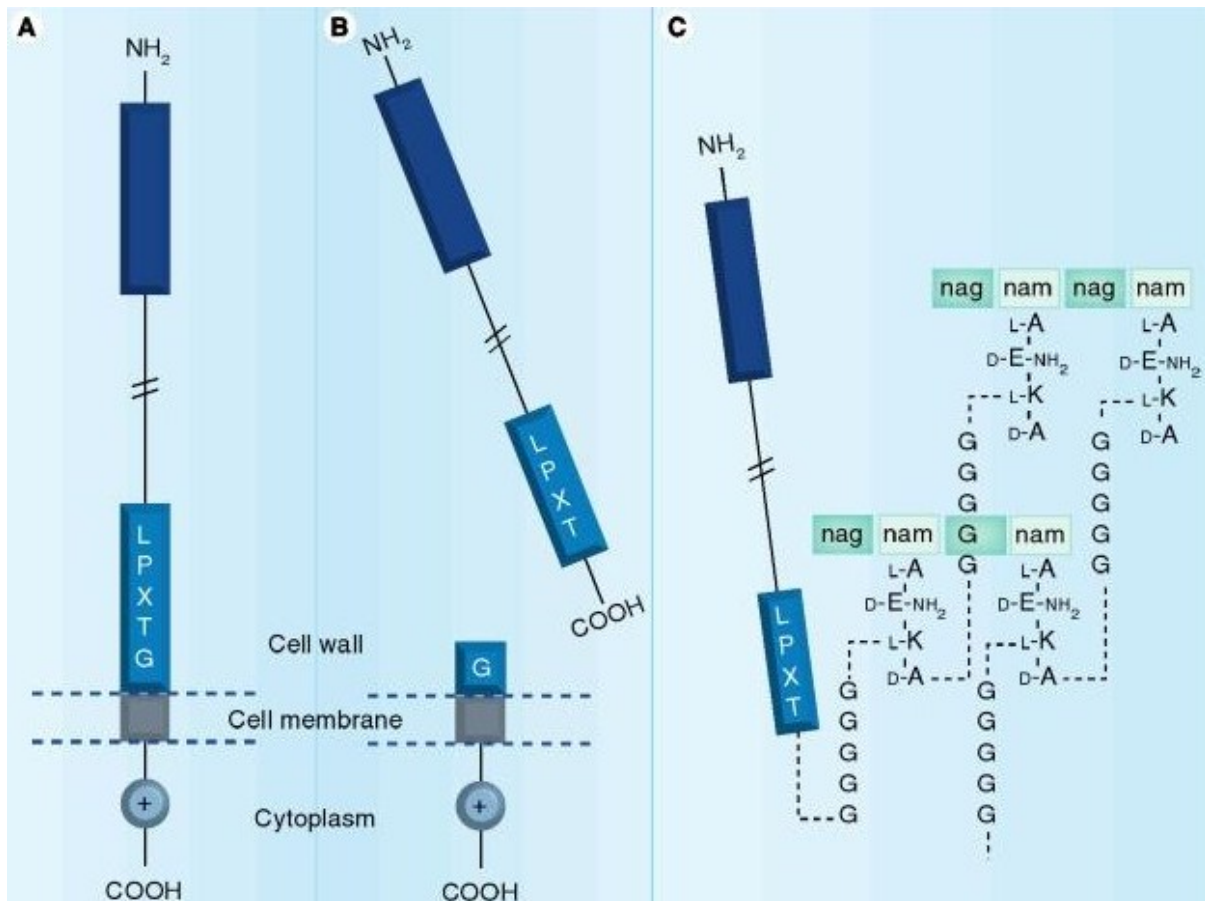


Figure 1.7: Example of the adhesion of sortase-dependent proteins to the Gram-positive bacterial cell wall (*S. aureus*).

(A) The sortase dependent protein with characteristic LPXTG motif, the hydrophobic domain is shown in grey. (B) The hydrophobic domain and the positively charged carboxyl-terminus prevent the complete secretion of the protein. The protein is instead cleaved between threonine and glycine residues of the LPXTG motif by the action of sortase enzyme. (C) The released amino-terminal portion of the protein becomes attached to the free amino-group of the peptidoglycan cross-linking pentaglycine bridge, anchoring the protein to the cell wall. Nag: *N*-acetylglucosamine; Nam: *N*-acetylmuramic acid (Neville 2010).

1.4.3. Non-covalent attachment of proteins to the Gam-positive surface

Although many Gram-positive surface proteins are attached covalently to stem peptides and thence peptidoglycan, some are attached via noncovalent ionic interactions to peptidoglycan or teichoic acids (Silhavy *et al.* 2010). For example, some proteins have been found to interact via C-terminal repeats with choline residues present in lipoteichoic or teichoic acids (Garcia *et al.* 1998). An example of this concerns the virulence factor, PspA in *Streptococcus pneumoniae* (Briles *et al.* 1998). A further method is by glycine and tryptophan repeats, referred to as 'GW modules' (Mesnage *et al.* 2000). GW modules are 80-amino acid tandem repeats starting with a GW di-peptide present at the protein C-terminus (Cossart and Jonquieres 2000) and initially identified in *L. monocytogenes* (Cabanès *et al.* 2002). The protein InlB of *L. monocytogenes*, a requirement for host cell invasion, was found to lack a LPXTG motif and instead have a series of GW repeats in its C-terminal region (Gaillard *et al.* 1991). Like PspA, InlB also contains an N-terminal Sec secretion signal and lipoteichoic acids are the cell wall ligand of GW modules (Jonquieres *et al.* 1999).

Some cell envelope proteins can also be membrane associated and anchored by a hydrophobic stretch of around 20 amino acids. This is preceded by a stop-transfer signal, a stretch of positively charged amino acids at the C-terminus of the protein. The actin-polymerizing protein, ActA of *L. monocytogenes* is an example of a membrane-anchored protein exposed to the medium (Kocks *et al.* 1992).

1.4.4. Lipoprotein biosynthesis

As originally described for Braun's lipoprotein of *E. coli*, proteins can also be localised to the bacterial cell envelope as lipoprotein (Braun *et al.* 1994). A specific signal peptide motif is required, known as a 'lipobox' for those proteins destined to become lipidated. The lipobox is characterised as having an L-3-[A/S/T]-2-[G/A]—1-C+1 sequence with a conserved cysteine as an absolute requirement (Hutchings *et al.* 2009; Rahman *et al.* 2008). Putative lipoproteins are typically directed initially through the Sec pathway, crossing the cytoplasmic membrane unfolded. Recently studies have also suggested a SecA2-dependent accessory Sec pathway (Gibbons *et al.* 2007; Lenz *et al.* 2003) and TAT pathways may also be used (Gralnick *et al.* 2006; Thompson *et al.* 2010; Widdick *et al.* 2011).

Biogenesis of Lipoproteins involves at least two enzymes: prolipoprotein diacylglyceryl transferase (Lgt) and lipoprotein signal peptidase (Lsp). Lgt initially catalyses the transfer of a diacylglyceryl moiety onto the conserved cysteine of the lipobox via a thioether linkage (Hutchings *et al.* 2009). Lsp then cleaves the signal peptide at the conserved cleavage site in the lipobox, ensuring the retention of the lipid-modified cysteine at the N-terminus of the mature lipoprotein. This pathway is conserved in all bacteria (Hutchings *et al.* 2009) whilst in Gram-negative bacteria there is a third step whereby, following Lsp mediated cleavage, Lipoprotein N-acyl transferase (Lnt) adds a further fatty acid in an amide linkage to the free amino group of the lipidated cysteine.

Homologues of Lnt have been identified in some high G/C content Gram Positive bacteria (Widdick *et al.* 2011) and there is evidence of N-acylation in some low G/C Gram-positive bacteria in the absence of Lnt homologues (Asanuma *et al.* 2011; Navarre *et al.* 1996). Some Gram Positive bacteria also possess more than a single copy of *lgt* with two

putative paralogues encoded in the genomes of *Bacillus cereus* ATCC10987 (Rasko *et al.* 2004), *Clostridium perfringens* (Shimizu *et al.* 2002) and *Streptomyces coelicolor* (Bentley *et al.* 2002; Thompson *et al.* 2010). The role of Lgt enzyme paralogues is currently unknown but has been suggested that a possible role could be the availability of an alternative lipoprotein specific pathway, akin to that of sortase paralogues. Bacterial lipoproteins have since been proposed to be functional equivalents of Gram-negative periplasmic proteins (Hutchings *et al.* 2009).

1.5. The molecular basis of virulence

Bacterial surface proteins have a wide variety of functions and are critically important in determining the success of a bacterial strain. Typical bacterial surface functions include those proteins involved in bacterial adherence or protection (either against toxins, environmental conditions or host responses). Some also provide bacteria with an enhanced ability to acquire nutrients, allow bacteria to interact with each other and compete for specific environmental niches. Some surface proteins are also essential for bacterial growth, maintenance and cell division.

General examples include MSCRAMM's which act as ligands; anti-phagocytic bacterial capsules; transporters to aid in the acquisition of nutrients; and extracellular toxins which damage or subvert the host immune system.

Understanding the role played by virulence factors can help elucidate the understanding of the pathogenesis of infection and identify potential points of treatment or vaccination (Mitchell and Mitchell 2010).

1.6. Virulence factors of *Streptococcus equi* subspecies and related pathogens

1.6.1. Fibrinogen-binding proteins

1.6.1.1. M-protein

Arguably the best studied cell wall-associated proteins of streptococci are the M protein family of *S. pyogenes* (Meehan *et al.* 1998; Smeesters *et al.* 2010) and other M-like proteins identified in Groups C and G streptococci (Bisno *et al.* 1997; Simpson *et al.* 1992).

The term M protein refers to fibrillar surface proteins of streptococcal species that have a variety of ligand activities. M proteins also have a variety of common structural attributes including variable N-terminal region, signal and cell wall anchoring (LPXTG) sequences, conserved central C repeats and C-terminal domains rich in proline. The term is also restricted to those proteins that induce opsonic antibodies for the strain from which the protein is derived. M protein has long established clear virulence functions primarily in inhibiting phagocytosis in the absence of opsonising antibodies (Smeesters *et al.* 2010). Strains that fail to express M protein are avirulent (Lancefield 1962). Key characteristics of M protein also include a hypervariable N-terminal region (Cunningham 2000) which results in a large degree of antigenic variation and the ability to elicit type-specific protective antibodies via the interaction of M protein with a wealth of host proteins (Smeesters *et al.* 2010).

S. zooepidemicus expresses two M-like proteins known as SzP (Moore and Bryans 1969) and SzM (Kelly *et al.* 2006) which are antigenically variable amongst strains and are

acid-extractable surface antigens. SzP elicits serum opsonic and protective responses and is a 40.1kDa, acid-resistant protein. It has also been shown to bind to human Hep-2 epithelial cells (Fan *et al.* 2008). Two similar M proteins are expressed in *S. equi*, one of which is highly sequence related to SzP and cross-reactive to anti-SzP antibodies, known as SzPSe (Timoney *et al.* 1995); the other is unique to strains of *S. equi* (SeM). Despite having functional similarities, SzP and SeM do not share a high degree of sequence homology to each other or indeed to the M proteins of *S. pyogenes* (Timoney *et al.* 1997). The M proteins of *S. pyogenes* have three distinct functional domains (A, B and C); regarding SzP or SeM only the C domain is present. They do however possess an N-terminal leader peptide and C-terminal cell wall sorting signal. SeM and SzPSe are also suggested to have extensive regions of α -helical coils as predicted by their secondary structures (Harrington *et al.* 2002). The M-like protein SzM of *S. zooepidemicus* is thought to be a replacement or modification of the SeM gene rather than a direct acquisition (Kelly *et al.* 2006).

Regarding SeM, sequence variation has been shown to be a key aspect *S. equi* virulence. Despite being uniform in size (~58 kDa), many isolates have been shown to possess a large degree of N-terminal variations (at amino acid positions 37-143) and occasionally signal peptide variation. For example, examination of a sample of horses with chronic guttural pouch empyema found up to 25% N-terminal deletion of the entire sequence had occurred (Chanter *et al.* 2000). N-terminal variance has also been noted in a variety of other isolates collected from 1950 onwards in North America, Japan and Europe (Anzai *et al.* 2005; Ivens *et al.* 2011; Kelly *et al.* 2006). Several other studies have shown that the rate of SeM mutation is high and evidence has been presented that changes can occur in the SeM allele in an infected horse during a outbreak (Anzai *et al.* 2005; Kelly *et al.* 2006). These studies also indicate that there may be diversifying selective pressure of variants with N-

terminal amino acid substitutions. Evidence also suggests that N-terminal variation affects only a conformational and not a linear epitope (Timoney *et al.* 2007). A possible explanation for the prevalence of N-terminal variation could be that SeM sequence changes enhance bacterial survival in the host. This conclusion is reinforced with emergent strains more commonly identified in chronically infected guttural pouches and cranial sinuses indicating that N-terminal variation may aid the process of immune evasion in these niches (Ijaz *et al.* 2011). Currently there are 87 allelic variants listed on the SeM database (<http://pubmlst.org/>).

Both SzP and SeM have been shown to bind to equine fibrinogen and immunoglobulin G (IgG) (Meehan *et al.* 2009; Timoney *et al.* 1997). The binding of fibrinogen reduces C3b deposition on the bacterial cell wall and resists neutrophil mediated phagocytosis (Boschwitz and Timoney 1994). This method of phagocytosis resistance is similar to that of the previously described M proteins of group A streptococci (Boschwitz and Timoney 1994). SeM antigens are also opsonised by antibodies against SeM that have no effect on SzP and vice-versa (Timoney *et al.* 1997). Studies have shown that resistance to phagocytosis in vitro in *S. equi* is predominantly due to the action of SeM with SzPSe having a more secondary role (Timoney *et al.* 1997). Investigations by Timoney and co-workers have also shown that the N-terminus of SeM contains epitopes reactive with mucosal IgA and stimulatory for lymph node T lymphocytes in immunized horses (Timoney *et al.* 2010). Variation of this region can also render the conformation of a B cell epitope un-reactive but surprisingly has no effect on fibrinogen binding or *S. equi* opsonisation.

1.6.1.2. H-binding protein

Recently a new fibrinogen binding protein unique to *S. equi* (Se18.9) was identified by Timmony and co-workers (Tiwari *et al.* 2007). This new protein was found to bind to complement regulatory protein factor H and found to reduce the bactericidal activity of equine neutrophils in serum. In the presence of Se18.9 specific antiserum, a two-fold increase in phagocytosis of *S. equi* was observed. As an anti-phagocytic protein, Se18.9 was shown to act a cofactor for factor I, inhibiting the C3 convertase pathway and thus reducing C3b deposition on the bacterial cell surface. Other proposed Se18.9 effects on opsonophagocytosis include direct toxic damage and neutrophil adhesion interference. Additionally since Se18.9 also binds to tonsillar epithelial cells and elicits strong serum and mucosal antibody responses it was theorised that Se18.9 may be directed by a similar acquired immune response selective pressure as illustrated by SeM (Ijaz *et al.* 2011).

In contrast however, to SeM it has since been shown that Se18.9 exhibits very little variation in a population of *seM* allelic variants. This suggests that Se18.9 structure may either be important for virulence function or may alternatively be minor compared to SeM (Ijaz *et al.* 2011). Although some *S. zooepidemicus* strains were shown to have similar sized open reading frames, no specific homologue has currently been found of Se18.9 in any known strains of *S. zooepidemicus*.

1.6.2. Hyaluronic acid capsule

In vertebrate species, hyaluronic acid (HA) is an important extracellular matrix polymer involved in a variety of functions such as cell trafficking and binding of water molecules. Investigations into pathogenic Group A and C streptococci have also found HA to be a key constituent of the streptococcal capsule material (Lee and Spicer 2000). Considering the morphology of *S. equi*, virulent isolates of *S. equi* are highly encapsulated producing mucoid colonies when cultured due to the presence of hyaluronic acid (Anzai *et al.* 1999). The capsule is a high molecular weight polymer of HA, a polysaccharide with alternating residues of N-acetylglucosamine and glucuronic acid (Timoney 2004). The capsule has several functions, the primary one being phagocyte resistance, preventing large numbers of streptococci from being associated with the surface of neutrophils. The negative charge of the capsule also produces a reducing environment, protecting the activity of streptolysin S and oxygen-labile proteases. The HA capsule of *S. pyogenes* has also been shown to act as a ligand for the receptor CD44 on human epithelial and haematopoietic cells (Harrington *et al.* 2002).

In *S. equi* an absence of the HA capsule has been shown to result in the aggregation of SeM, thereby causing a loss of the three-dimensional conformation and preventing the functionality of SeM (Timoney, 1999). The adherence of *S. zooepidemicus* to HeLa cells has also been shown to be significantly reduced in non-encapsulated strains (Wibawan *et al.* 1999).

Synthesis of the HA capsule is by the multifunction membrane-bound enzyme, HA synthase which assembles HA from two precursor sugars, UDP-glucuronic acid and UDP-N-acetylglucosamine (Tlapak-Simmons *et al.* 1999).

The *has* operon gene products are responsible for the synthesis of the capsule. The operon consists of (the aforementioned) hyaluronate synthase (*hasA*) (Deangelis *et al.* 1993), UDP-glucose dehydrogenase (*hasB*) and UDP-glucose pyrophosphorylase (*hasC*) (Dougherty and Vanderijn 1994). Two *has* operons were recently identified in *S. equi* and *S. zooepidemicus* and are likely the result of intragenome gene duplications (Blank *et al.* 2008).

In *S. zooepidemicus* (H70), *hasC* has been translocated to an opposite replicore by a large genome inversion (Blank *et al.* 2008). The constituents of the *has* operon in *S. zooepidemicus*; *hasA*, *hasB* and *hasC*, as well as N-acetyl-glucosamine-1 phosphate uridyltransferase [*glum*] and Glucose-6-phosphate isomerase [*pgi*] are all located in order (Blank, *et al.* 2005). However, in *S. equi* the genes are separated onto two operons; *hasA*, *hasB*, *hasC* and *hasC*, *glmU* and *pgi* (Blank *et al.* 2008). Both Pgi and GlmU are both involved in the synthesis of the UDP-N-acetyl-glucosamine precursor, with the *hasB* and *hasC* gene products responsible for the synthesising UDP-glucuronic acid precursor.

The *hasC*-mediated inversion in *S. equi* (4047) has been theorised to account for an increased production of hyaluronate capsule compared to isolates of *S. zooepidemicus* (Holden *et al.* 2009). The inclusion of additional genes in the *has* operons of both *S. equi* and *S. zooepidemicus* also indicates a selective advantage of enhanced HA production with *S. zooepidemicus* able to direct more than 10% of its carbon towards HA production, far higher than other streptococci (Blank *et al.* 2005).

1.6.3. Lipoproteins

Currently five lipoproteins have been identified in *S. equi* and *S. zooepidemicus*; hyaluronate-associated protein (HAP) (Chanter *et al.* 1999), metal binding lipoprotein (MBL) (Harrington, *et al.* 2000), SeHtsA (Nygaard *et al.* 2006), PrtM (Hamilton *et al.* 2006) and the acid phosphatase lipoprotein, LppC (Hamilton *et al.* 2000).

HAP was first described in *S. equisimilis* as a hyaluronate synthase (Lansing *et al.* 1993) and is a 56-kDa extracellular threonine kinase thought to relate to the regulation of HA capsule synthesis and HA release from the cell surface (Lansing *et al.* 1993). This is said to have been achieved through HAP autophosphorylation (Nickel *et al.* 1998). Two proteins similar to HAP were subsequently identified in *S. equi* and *S. zooepidemicus* and were thought to be hyaluronate (capsule)-associated proteins (Chanter *et al.* 1999). It has since been shown that they are instead likely to act as substrate-binding lipoprotein for the uptake of oligopeptides via ATP-binding cassette (ABC) transporters (Harrington *et al.* 2002).

The ability of bacteria to scavenge a wide variety of nutrient substrates is a key aspect of infection. Bacterial ABC import transporters are typically ATP-driven three component systems involving a substrate-binding protein (SBP), a trans-membrane transport unit and a cytoplasmic ATP-binding component (Facey and Kuhn 2010). The majority of ABC transporters are importers and a characteristic component of such systems is a solute binding lipoprotein encoded within the same operon as the other previously stated ABC transporter components (Harrington *et al.* 2000).

A well characterised SBP is PsaA, a member of the LraI family of metal-binding lipoproteins and a substrate-binding component of a manganese import system, identified in

S. pneumoniae (Johnston *et al.* 2004). PsaA has been shown to perform a role in bacterial protection from superoxide with mutant strains of *psaA* shown to cause deficiencies in growth, virulence, adherence and the oxidative stress response (Johnston *et al.* 2004).

The LraI family of proteins also consists of MtsA of *S. pyogenes* and SloC of *Streptococcus mutans*. Mutant strains with an absence of SloC have been shown to possess a decreased virulence in animal model studies (Kitten *et al.* 2000). *S. equi* contains a homologue of PsaA (87% similarity) and, MtsA (95% similarity), originally termed MBL (Harrington *et al.* 2000). Despite the PsaA homologue also being present in *S. zooepidemicus*, western blot analysis using anti-PsaA serum and whole cell extracts showed a weaker level of detection in *S. zooepidemicus* than that of *S. equi*. A plausible explanation for this occurrence is that gene expression of this SBL is lower in *S. zooepidemicus* during standard growth conditions; however the employment of different bacterial growth conditions did not intensify the reaction (Harrington *et al.* 2000) although this could also be due to weaker cross-reactivity.

The acquisition of iron is notably essential to establish infection (Wooldridge and Williams 1993). In Gram-positive pathogens, this is usually achieved via heme specific ABC transporter systems and cell surface proteins capable of binding heme, known as hemoproteins. Considering *S. pyogenes*, a heme-specific ABC transporter is encoded by 8 specific genes and regulated by MtsR. One such operon *htsABC*, with a lipoprotein homologue demonstrated in *S. equi* notated as SeHts. Studies have shown that SeHtsA binds heme and hemin and consequently this iron acquisition mechanism is likely to be conserved between *S. equi* and *S. pyogenes*. In addition to this system, a siderophore type ('equibactin') iron-uptake locus has also been identified in *S. equi* (Heather *et al.* 2008).

A specific lipoprotein known as maturase (PrtM) has also been characterized as a virulence factor. Mutants of PrtM were shown to have an impaired ability to colonize the equine upper respiratory tract (Hamilton *et al.* 2006). In addition to these 5 lipoproteins, genomic analysis of *S. equi* and *S. zooepidemicus* suggests that at least 30 other putative lipoproteins are encoded in these genomes (Sutcliffe, unpublished data). The acid phosphatase lipoprotein, LppC will be discussed in greater detail in section 1.19 and is the subject of Chapter 5 of this thesis.

1.6.4. Superoxide dismutase

S. equi has been shown to encode a manganese-dependent superoxide dismutase (SOD) (Poyart *et al.* 1998). SOD catalyses the conversion of superoxide anions to hydrogen peroxide and water, converting reactive oxygen intermediates into less toxic forms. The generation of reactive oxygen intermediates is associated with phagocytosis and thus SOD may allow *S. equi* to combat oxidative stress. Although not investigated in *S. equi* it has been shown that inactivation of the manganese-dependent SOD gene (*sodA*) of *S. pneumoniae* led to a reduction in pneumococcal virulence in a mouse model (Gibson and Caparon, 1996). The gene *sodA* has been shown to be conserved amongst all *S. equi* isolates investigated (Poyart, *et al.* 1998) and has become a target gene of interest for real-time PCR based detection and differentiation of *S. equi* and *S. zooepidemicus* (Båverud *et al.* 2007).

1.6.5. Super-antigens

Emerging strains of *S. pyogenes* with particularly high mortality and morbidity were identified in the 1980s to produce super-antigens (sAgs) (Ikebe *et al.* 2002). sAgs interfere with the development of a protective immune response, by disrupting antigen-specific T cell responses and inhibiting the production of antibodies (Llewelyn and Cohen 2002). As well as disrupting the mechanism of major histocompatibility complex (MHC)- restricted antigen presentation (Dellabona *et al.* 1990), the ability to augment endotoxin activity and cause local inflammation are other related properties of sAgs (Llewelyn and Cohen 2002). The stimulation of a non-specific T cell response and proinflammatory cytokine release results in toxic shock syndrome in individuals suffering a *S. pyogenes* infection (Artiushin *et al.* 2002; Sriskandan *et al.* 2007).

Currently 11 sAgs have been identified in *S. pyogenes* with some chromosomally located and others located on mobile genetic elements (Smoot *et al.* 2002; Sriskandan *et al.* 2007). Several clinical features of strangles disease are associated with proinflammatory cytokine release and therefore *S. equi* sAgs are of great interest (Artiushin *et al.* 2002). Four homologues of *S. pyogenes* sAgs have been identified in *S. equi*: *seeH*, *seeI*, *seeL*, and *seeM* (Holden *et al.* 2009). All *S. equi* sAgs are carried on two prophages i.e. Φ Seq3 (*SeeM* and *SeeL*) and Φ Seq4 (*SeeH* and *SeeI*) with *SeeL* and *SeeM* identified in *S. equi* strains that predate the isolation of sAg encoding *S. pyogenes* strains (Ikebe *et al.* 2002). The pyrogenicity of *SeeH* and *SeeI* were recently investigated and found to result in a raised temperature in animal models post injection (Artiushin *et al.* 2002). *SeeH* and *SeeI* also appeared to be potent for equine peripheral blood mononuclear cells, at least when compared to *Spe-A* and *Spe-C* for human T cells (Braun *et al.* 1993). Recently *S. equi* sAg-induced lymphoproliferation was also shown *in vitro*, with *SeeM* indicating the highest degree of potency. Deletion of sAg genes was also shown to annul mitogenic activity (Paillot *et al.*

2010). It has also been speculated that differences in the level of sAgs expression amongst strains may account for strain specific differences in disease severity, although this has yet to be clarified (Paillot *et al.* 2010).

It was previously thought that sAgs were absent in *S. zooepidemicus* as strains with mitogenic activity were screened for the presence of *S. equi* sAgs homologues and none were found (Holden *et al.* 2009). However three novel sAgs - SzeF, SzeN, and SzeP from *S. zooepidemicus* (strain BHS5) have been recently identified and were found to be significantly associated with isolation from equine cases of non-strangles lymph node abscessation (Paillot *et al.* 2010).

1.6.6. Fibronectin-binding proteins

Fibronectin (Fn) is found in plasma and in the extracellular matrix in a fibrillar form and is a dimeric glycoprotein (Lindmark *et al.* 1999). Specific cell surface receptors bind to specific domains of fibronectin, mediating the substrate adhesion of eukaryotic cells. Fibronectin also interacts with a variety of macromolecules including collagen, fibrin, heparin and DNA (Lindmark *et al.* 1999).

Fn-binding proteins are cell surface proteins which have been shown to enhance bacterial potential to cause disease (Lindmark *et al.* 2001; Schwarz-Linek *et al.* 2003). The Fn-binding proteins SfbI/protein F1 and M1 from *S. pyogenes* have been shown to enable the spread of bacteria deeper into tissues (Jadoun *et al.* 1997) by mediating the adherence to and the invasion of epithelial cells (Lapenta *et al.* 1994). Fn-binding proteins have also been shown to mediate host cell invasion without the presence of other virulence factors in *S.*

aureus (Sinha *et al.* 1999) and *S. pyogenes* (Ozeri *et al.* 1998). Fn-binding proteins may also be significant in the pathogenesis of streptococcal endocarditis (Jakubovics *et al.* 2009).

Currently three Fn-binding cell proteins, FNZ (Lindmark, *et al.* 1996), FNZ2 (Hong 2005) and secreted fibronectin binding protein (SFS) (Lindmark and Guss 1999) have been identified in *S. zooepidemicus*. FNZ and FNZ2 has been shown to encode all the features typically associated with cell wall associated proteins including a C-terminal domain containing LPXTG, a hydrophobic spanning domain and a signal peptide (Lindmark *et al.* 1996).

Homologues of FNZ and FNZ2 have also been identified in *S. equi*, notated as FNE (*seq0375*) (Lindmark *et al.* 2001) and FNEB (Lannergard *et al.* 2005). SFS has also been identified in all investigated strains of *S. equi*. FNE however binds fibronectin considerably more weakly than FNZ due to a one-base deletion present in the *fne* gene, enabling the translation of only the 5'-terminal portion of the gene and thus an absence of terminal sequence motifs involved in protein cell wall anchoring. FNZ also possess two Fn-binding sites in the C-terminal and N-terminal region in contrast to FNE (Lannergard *et al.* 2005). SFS was found to inhibit binding between collagen and Fn (Lindmark and Guss 1999). However studies have shown that SFS has a limited effect on Fn binding in *S. equi* and it is not known whether or not SFS is cell bound or secreted (Lindmark and Guss 1999).

Horses with strangles have been shown to produce a strong IgG antibody response to FNE, SFS and FNEB. However antibody titres to each have no correlation, indicating that while an antibody response is generated to each individual Fn-binding protein, there is no general Fn-binding antibody response, with each individual horse responding differently to different antigens (Lannergard *et al.* 2005).

1.6.7. Collagen-binding proteins

The ability of bacteria to adhere to host collagen via collagen-binding protein has been characterised as a virulence factor in *S. aureus* (Hienz *et al.* 1996). A collagen-binding protein, similar to CNA of *S. aureus* (231/657 amino acids), was identified in *S. equi* and denoted as CNE (Lannergard *et al.* 2003). The highest degree of homology between CNE and CNA was also found to be in the collagen binding region.

The *cne* gene encodes a 657 amino acid protein and contains an N-terminal signal sequence, a hydrophobic C-terminal transmembrane region and an LPXTG motif (LPDTG). The *cne* gene has also been identified in *S. zooepidemicus* but is not identical to that of *S. equi*. Although the significance of CNE in virulence has not been established, an evolutionary relationship between CNE and CNA is suggested (Lannergard *et al.* 2003). Following a study into putative extracellular proteins in *S. equi* using signal sequence phage display, another collagen-like protein was identified and notated as SclC (Karlstrom *et al.* 2004). The identified protein was found to share similarity to the N-terminus of SclB, a collagen-like protein of *S. pyogenes* (Whatmore 2001). However no specific binding activity of SclC has currently been shown and despite high collagen-binding activities in *S. equi* and *S. zooepidemicus* no collagen-binding activity was detected in material released from the cell surface (Karlstrom *et al.* 2004). This is in agreement to similar findings concerning CNE (Lannergard *et al.* 2003).

1.6.8. Immunoglobulin G (IgG) endopeptidases

An important participant of the adaptive immune system is IgG, with the binding of IgG to an antigen providing an enhancement to several reactions including complement activation and phagocytosis (Hulting *et al.* 2009). In order to evade recognition by immunoglobulins, some bacterial pathogens possess IdeS which is a cysteine protease that cleaves IgG and therefore interferes with phagocytic killing. An example of this is IdeS from *S. pyogenes* (von Pawel-Rammingen *et al.* 2002). A homologue of IdeS has since been found in *S. equi* and *S. zooepidemicus*, notated as IdeE and IdeZ respectively (Lannergard and Guss 2006). Cell culture studies show that active IdeE is expressed during the logarithmic phase only in the presence of horse serum (Lannergard and Guss 2006). IdeE has also been shown to block phagocytosis *in vitro* by reducing the bactericidal activity of equine neutrophils (Timoney *et al.* 2008).

Recently a further IgG endopeptidase was found in *S. equi*, with a similar protein also present in *S. zooepidemicus*, notated as IdeE2 and IdeZ2 respectively (Hulting *et al.* 2009). Sequence comparisons have shown that IdeE2 and IdeZ have a closer relationship to IdeS than IdeE/IdeZ2 and that IdeE and IdeE2 were conserved amongst all *S. equi* strains investigated. In contrast, IdeZ and IdeZ2 were only found in a selection of *S. zooepidemicus* strains (Hulting *et al.* 2009). IdeE and IdeZ were also found to exhibit a broader substrate activity than either IdeE2 or IdeZ2. Immunization experiments have also demonstrated that IdeE and IdeE2 induce protection in a mouse model with IdeE2 in particular shown to significantly reduce both nasal growth of *S. equi* and improve the condition of vaccinated animals (Hulting *et al.* 2009).

1.7. Vaccines against *S. equi*

Conventional vaccination strategies have proved largely ineffective towards *S. equi*. Little efficacy and an undesirable level of adverse reactions were shown when trialling vaccines containing inactivated whole bacteria or extracts (Lr 1990). Initially a protein-rich acid extract (Reif 1981) and a mutanolysin extract (Bryant *et al.* 1985) vaccine was introduced in the 1980s in North America but showed only a moderate level of efficacy with no reduction in bacterial shedding or a reduced transmission. Vaccinations based on bacterins and adjuvanted extracts were also shown to have little efficacy (Jorm 1990). A vaccine consisting of a non-encapsulated strain of *S. equi* was then later developed and used (Pinnacle INtm) as a nasal vaccine but never licensed in Europe. A second live attenuated vaccine (Equilis StrepE) was also developed and granted a licence in Europe (Jacobs *et al.* 2000), but later removed from use due to safety concerns (Kemp-Symonds *et al.* 2007; Newton *et al.* 2005). Equilis StrepE has since been re-launched in Europe, initially in the Netherlands in May 2010. However a need for an improved vaccine remains prominent.

Further vaccination strategies are based on ‘reverse vaccinology’, using genome sequence data of *S. equi* to identify surface exposed and secreted proteins. The target protein initially of many of the early vaccines using this strategy was the SeM protein. Often vaccination challenge studies using the mouse model were encouraging. However, despite the presence of SeM reactive antibodies, significant protection in horses was not shown (Sheoran *et al.* 2002) and high level of side effects with SeM based vaccination trials had previously been shown (Hoffman *et al.* 1991). The general consensus for an improved vaccination strategy, based on the shortcomings of past vaccine trials are that of multiple protective

epitopes. An alternative to live attenuated vaccines are therefore multi-component subunit based vaccines of several immunogenic *S. equi* surface proteins (Waller and Jolley 2007b).

The first vaccine based on combinations of subunits contained the surface proteins FNZ, SFS and EAG. The use of EAG was shown to demonstrate protection as a single antigen but its efficacy was significantly improved by the addition of FNZ and SFS (Flock *et al.* 2004).

Using recombinant antigens against *S. equi*, numerous groups have shown some success in mouse model systems, similar to the SeM based vaccine. Partial protection was achieved in ponies challenged with ('Trivacc') (Waller *et al.* 2007) consisting of EAG (Flock *et al.* 2004), SclC (Karlstrom *et al.* 2004) and CNE (Lannergard *et al.* 2003). However, despite promising reductions in nasal discharge, bacterial carriage and empyema the level of protection in individual horses was limited (Waller *et al.* 2007). The 'Trivacc' vaccine has since been improved and revaluated ('Septavacc'). Two cell surface anchored proteins of unknown function (SEQ0256 and SEQ0402) were added to EAG, CNE and SclC as were the two endopeptidases (IdeE and IdeE2), which resulted in an improved level of protection with only one pony out of seven developing lymph node abscesses (Guss *et al.* 2009).

A further recent development has concerned the cell envelope proteinase, SpyCEP (*cepA*). SpyCEP was found in *S. pyogenes* to represent a convincing vaccine candidate due to its high degree of conservation between *S. pyogenes* isolates (Turner *et al.* 2009). Data showed that immunity to SpyCEP prevented severe infection dissemination in eight BALB/c mice. The group also investigated functional homologues of SpyCEP in other streptococcal species with homologues of 74% similarity to SpyCEP identified in *S. equi* (SeCEP) and *S. zooepidemicus* (SzoCEP). Cross-protectivity of SpyCEP immunization was also investigated and showed that immunization against SpyCEP induces neutralizing immunity, preventing

dissemination of disease from the site of infection (Turner *et al.* 2009). However *S. equi* is more suppurative than *S. pyogenes* and despite a reduction of disease dissemination the infection was not completely inhibited. Little protection was shown with an immunization challenge against *S. equi* and SzoCEP (98% identical to SeCEP) was found to be poorly expressed in the strains tested.

Common difficulties with many of the recent vaccine attempts include a short duration of immunity and purulent reactions at the site of injection. The emphasis on subunit and inactivated vaccines has also created a need for more competent adjuvants and delivery systems (Meeusen *et al.* 2007). An example of how to address this need is the use of surface modified polymeric nanoparticles, which displayed an improved adsorption efficacy for *S. equi* extract proteins (Florindo *et al.* 2010).

1.8. Phospholipases

Phospholipids were once thought of as inert building blocks of cellular membranes, but are now considered important metabolic entities with phospholipids and their metabolites providing prominent roles in many cellular control systems. The elevation of the importance of phospholipids has thus given phospholipases (PL) a preeminent standing with regard to potential roles of phospholipases in biological systems. PL are now regarded as more than simple maintenance enzymes being considered as important metabolic entities (Dennis 2000) and have now been shown to play a role in a large variety of biological functions, including; host defence, signal transduction and lipid mediator production (Lambeau and Lazdunski 1999), along with the ability to destroy or disrupt components of host cells membranes (Ghannoum 2000).

The PL comprise of an expanding family of distinct enzymes which share the ability to hydrolyze one or more ester linkage in glycerophospholipids to release fatty acids and lysophospholipids (Lambeau and Lazdunski 1999). Differentiation of PL into groups PLA, PLB, PLC and PLD indicate the specific bond targeted in the phospholipid molecule (Figure 1.8).

The repertoire of potential functions and impact on host biological functions ensure that characterised PL enzymes are of great interest regarding the search for novel virulence determinants of pathogenic bacteria and two recently identified PL enzymes in *S. equi* are the subject of chapter 4.

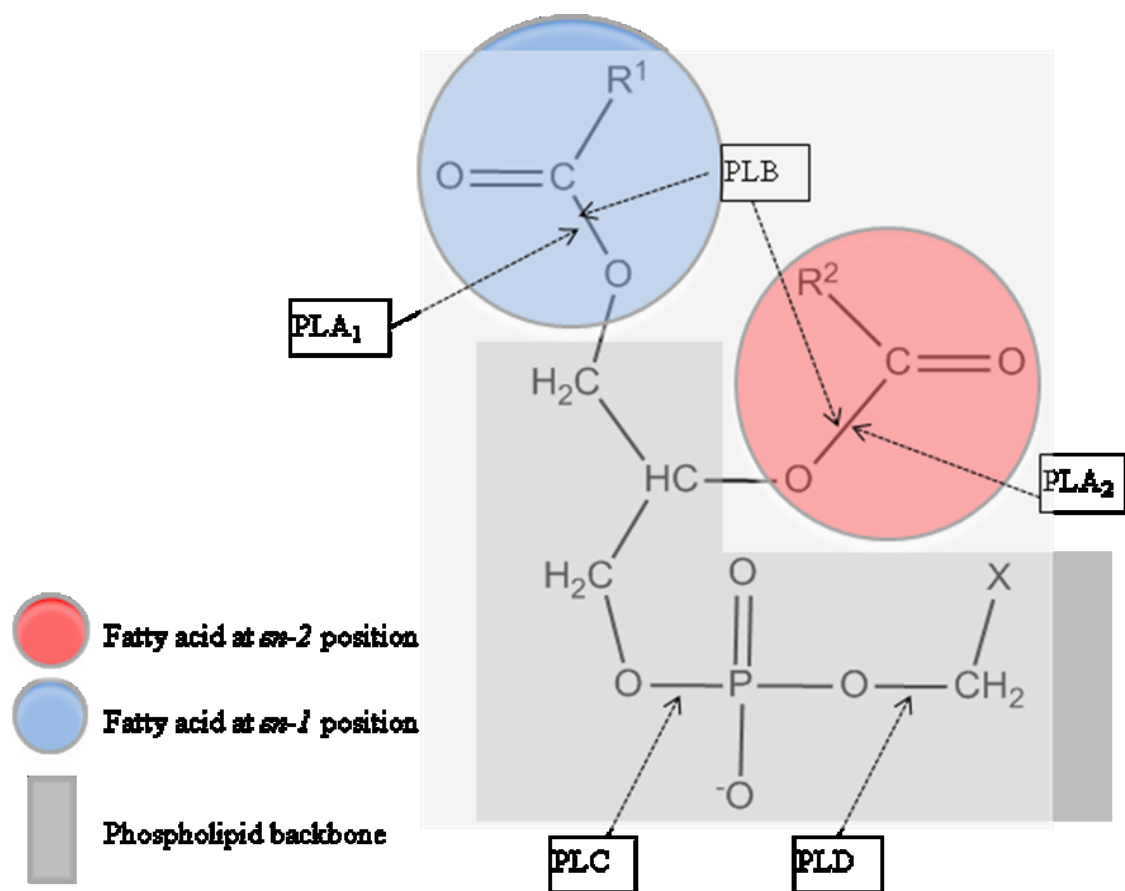


Figure 1.8: Current classification of phospholipases.

Group A cleave the fatty acid carboxylic acid ester from phospholipids. Group A are further divided into two sub-groups (PLA₁ and PLA₂) based on position specificity of fatty acid cleavage; either *sn*-1 (PLA₁) or *sn*-2 (PLA₂). R¹ and R² represent fatty acid chains in either the *sn*1 or *sn*2 position of the phospholipid, respectively. PLB can result in total deacylation of phospholipids with the potential of dual PLA₁ and PLA₂ activity. PLC and PLD catalyze the hydrolysis of one of the two phosphate ester bonds. X refers to the phospholipid head group.

1.8.1. Phospholipase putative roles in bacterial virulence

Historically PL enzymes have been shown to have a well-defined function regarding the damage and destruction of biological tissue. A large body of literature is available regarding venom secreted phospholipase enzymes, which can be potent toxins exerting a wide range of effects including neurotoxicity, myotoxicity, cardiotoxicity and inflammation (Lambeau and Lazdunski 1999).

The number of identified and classified PL enzymes from bacterial and protozoan pathogens has recently been increasing at a steady rate. Although many PLs have now been identified in bacteria, many have as yet undefined roles. Although originally thought of as possessing sole putative roles in membrane maintenance, mutants have shown that some PLs are not necessarily essential for viability (Oki *et al.* 1972). In view of this, several PL enzymes have now been identified as important virulence factors for pathogenic bacterial species. A common theme is the breakdown of host cell barriers and the creation of a more favourable bacterial niche.

Most characterised PL virulence factors belong to the PLC class. Indeed, historically, initial investigations regarding phospholipases as virulence factors concerned the reclassification of some prominent bacterial toxins such as α toxin of *Clostridium perfringens* (Macfarlane 1941) and β toxin of *S. aureus* (Doery *et al.* 1963) as PLC. PLs as toxins are generally linked to the active destruction of membrane phospholipids (cytolysis), aiding the bacterium to degrade and penetrate the mucous layer or to avoid immune cells. Other non-cytolytic tissue destruction has also been attributed to the presence of PLC from *C. perfringens*. Incubation with purified α -toxin also induced the expression of cytokines including tumour necrosis factor (TNF- α), platelet activating factor (PAF) (Stevens *et al.*

1997) and Interleukin 8 (IL-8) (Bryant and Stevens 1996) as well as the triggering of the arachidonic acid cascade (see section 1.8.2). In contrast some PLs have more subtle roles in bacterial pathogenicity. The facultative intracellular pathogen *L. monocytogenes* possesses two PLC enzymes, PlcA and PlcB. Mutation of PlcA demonstrated that Plc provides a clear role in bacterial escape from the macrophage phagosome (Camilli *et al.* 1993). Further to the evidence provided from investigations into PLCs, PLD secreted by *Corynebacterium pseudotuberculosis* has been shown to play a major role in caseous lymphadenitis, enhancing the spread of infection through tissues to the lymphatics by increasing vascular permeability as well as being dermonecrotic (Muckle and Gyles 1983).

Recently more characterised examples of the lesser understood PLA₂s have begun to emerge. For example, *Yersinia enterocolitica* secretes PLA₂ which has been shown to contribute to enhanced bacterial colonisation in a mouse model (Schmiel *et al.* 1998). A reduced bacterial colonization was exhibited in the Peyer's patches and mesenteric lymph nodes of mice infected with mutants of *Y. enterocolitica* lacking PLA₂.

Bacterial haemolytic and cytotoxic activities have also been attributed to the activity of PLA₂. For example bacteria belonging to the genus *Rickettsia* have been shown to display haemolytic activity, requiring a mildly acidic pH and divalent cations consistent with PLA₂ activity. The role of rickettsial PLA₂ was confirmed by showing effective inhibition of the cytotoxic effect of *Rickettsia prowazekii* following the addition of the PLA₂ inhibitor bromophenacyl bromide (BpB) and by western blot studies with anti-phospholipase A antibody from king cobra venom (Walker *et al.* 2001).

Several PLA₂ enzymes have also been implicated as virulence factors in *Helicobacter pylori*. The role acts by hydrolysing the fatty acid from the second carbon of membrane phospholipids such as lysolecithin, the metabolism of lysolecithin therefore occurs slower

than fatty acids, allowing the bacteria to persist in host tissues longer (Langton and Cesareo 1992). Recently a PLA₂ (PlaB) was characterised as a new virulence factor in *Legionella pneumophila*. This cell-associated PLA was found to express contact-dependent haemolytic activity and mutant strains were identified to have an impaired replication in the lungs and dissemination to the spleen in a guinea pig infection model (Schunder *et al.* 2010).

A further well characterised family of prokaryotic PL is the OMPLA (outer membrane phospholipase A) family found in many Gram-negative bacteria. Originally identified in *E.coli* (Scandell and Kornberg 1971), several similar OMPLA have been identified in *H. pylori* and several enterobacteriaceae. Regarding *H. pylori*, a specific characteristic of this human pathogen is its ability to reversibly alter its membrane lipid composition from <2% lysophospholipid (a product of phospholipase A₂ reactions) in one variant to >50% in another (Tannaes *et al.* 2000). This increase of lysophospholipids in *H. pylori* is due to action of OMPLA on the ordinary phospholipids, rather than the lipopolysaccharide anchor. This then increases the bacteria's tolerance to pH and improves host cell attachment. This increase in lysophospholipid allows the release of further *H. pylori* virulence factors such as VacA and urease, overall altering the host environment to favour bacterial growth, colonization and survival (Istivan and Coloe 2006).

When the fatty acid, arachidonic acid (AA) is released from phospholipids, PLA also provide key precursors for the generation of eicosanoids by several eicosanoid-generating enzymes (See section 1.8.2). AA metabolism has also been shown to regulate *E. coli* penetration of the Blood-Brain barrier (Zhu *et al.* 2010). Although the exact mechanisms remain unknown, gene deletion of the PLA₂ of *E. coli* K1 and pharmacological inhibition resulted in reduced bacterial penetration into the brain, but not into non-brain organs (Das *et al.* 2001)

1.8.2. The mammalian arachidonic acid (AA) cascade

AA (cis-, cis-, cis-, cis-5, 8, 11, 14-eicosatetraenoic acid) is a 20-carbon polyunsaturated fatty acid (C20:4) and axial eicosanoid precursor derived from dietary sources (linoleic acid) in mammals once AA is biosynthesised, it is esterified into the phospholipids of the other cell membranes (Hyde and Missailidis 2009).

AA and its metabolites are therefore essential fatty acids in the majority of mammals. AA released by PLA₂ is the key constituent of the AA cascade (Fig 1.9). The AA metabolites produced during the AA cascade are termed eicosanoids and produced by the catalytic activities the constituents of three pathways: cyclooxygenase (COX), lipoxygenase (LOX) and cytochrome P450 epoxygenase (CYP450) (Hyde and Missailidis 2009). Eicosanoids include lipid signal mediators which play a key role in cellular signalling cascades. The majority of AA metabolites can act as both pro- and anti- inflammatory mediators by modulating gene expression, cytokine signalling and other regulatory factors (Cabral, 2005).

Drugs with the action of inactivating human PLA₂ and thus inhibiting the AA cascade are currently of much pharmacological interest regarding potential suitability as anti-inflammatory targets (Sugiyama *et al.* 2002).

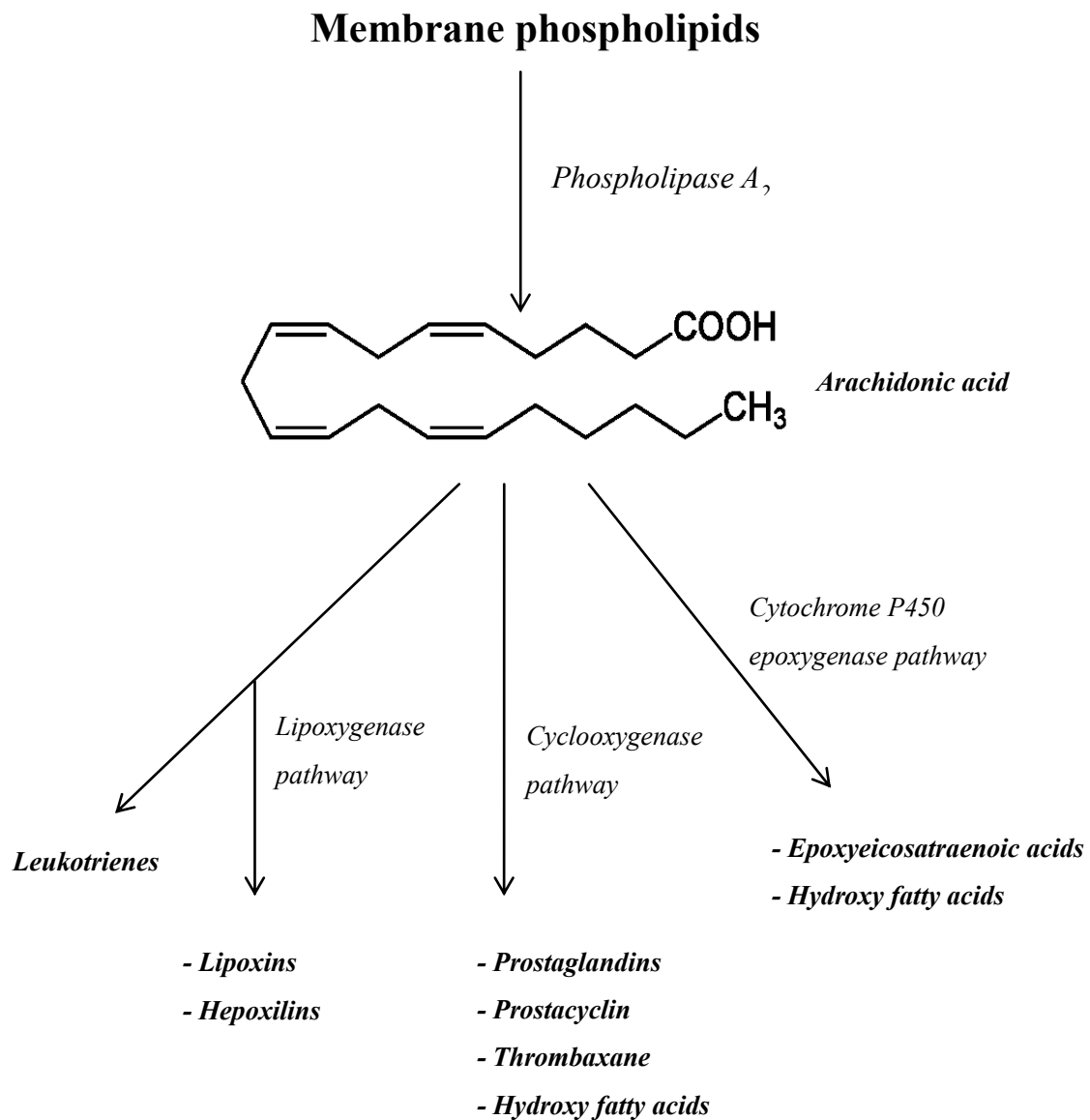


Figure 1.9: The AA cascade, indicating the three key enzymatic pathways of AA.

Taken from (Hyde and Missailidis 2009)

1.8.3. PLA₂

As described in section 1.8.1, the PLA₂ group of PL enzymes are of particular interest as potential virulence factors. This group of enzymes can be either high molecular weight cytosolic (CPLA₂) or low molecular weight secreted (sPLA₂) and represent a super family which also contains calcium independent PLA₂ (iPLA₂), the latter being widely distributed in eukaryotic cells and tissues (Madsen, et al. 2011).

sPLA₂ have also been further classified into currently 11 distinct groups (I-XI) and a variety of sub-groups, based upon their amino acid sequences. At present, each group has at least one member with a majority containing two or more members (Meyer, et al. 2005). All the above groups of sPLAs share a common catalytic mechanism of activity. The active site is similar to that of the Ser/His/Asp catalytic triad found in proteases (Poi, et al. 2003) and calcium ions are a requirement for sPLA₂ and provide coordination between the substrate and the main carbonyl oxygen (aspartate).

A water molecule acts as the nucleophile in the attack on the *sn*-2 position of the carbonyl carbon in the phospholipid, acting in the same role as the serine in proteases (Madsen, et al. 2011). A tetrahedral intermediate is thus formed and the *sn*-2 oxygen becomes protonated, releasing the products (a fatty acid and a lysophospholipid).

All sPLA₂ enzymes characterised so far are water-soluble and are required to bind to organized lipid interfaces for catalytic turnover. Naturally occurring phospholipids have poor solubility in water as monomers. Therefore it is necessary for hydrolysis to occur at the lipid-water interface (Ghomashchi, et al. 1998). All phospholipases are subjected to an increase in

activity when the lipid concentration is above the critical micelle concentration (interfacial activation) (Sekar, et al. 1997).

Different subtypes of sPLA₂ exhibit lipid substrate specificity, for example both subgroups V and X are able to catalyze the hydrolysis of both anionic and zwitterionic phospholipids in contrast with subgroup II, which is only active towards anionic phospholipids. CPLA₂ differs in catalytic mechanism, as the presence of Ca²⁺ is not required at the active site. These enzymes are translocated from the cytosol to intracellular membranes in response to an increased intracellular Ca²⁺ presence (Schievella, et al. 1995).

1.8.4. sPLA₂ of streptococci species

A prophage-encoded extracellular PLA₂, designated SlaA of *S. pyogenes* has been recently identified (Nagiec, et al. 2004). SlaA expression in *S. pyogenes* (strain MGAS315) has been shown to be induced when exposed to human pharyngeal epithelial cells indicating a role in virulence (Banks, et al. 2003). It has been recently identified that the acquisition by a *S. pyogenes* strain of an SlaA encoding bacteriophage has been responsible for the vast majority of human infections by serotype M3 strains in many countries (Sitkiewicz, et al. 2006). SlaA contains a region of conserved C-terminal amino acid residues; homologous to several sPLA₂ enzymes including toxins produced by venomous snakes such as textilotoxin produced by the Australian common brown snake, *Pseudonaja textilis*. SlaA has since been shown to have enzymatic activity against phospholipids with several head groups, releasing the acyl chains in the *sn*-2 position including the ability to release AA (Nagiec, et al. 2004). Evidence regarding SlaA as a virulence factor of *S. pyogenes* has been established using isogenic

mutant strains with genetically inactivated *slaA* (Sitkiewicz, et al. 2006). Purified recombinant SlaA was found to not express cytotoxic activity to cultured epithelial cells when added exogenously but the mutant lacking SlaA was found to be significantly less toxic to epithelial cells.

It has been suggested that SlaA enters host epithelial cells through an active transport mechanism utilising an as yet unknown host-cell receptor (Sitkiewicz, et al. 2007). Entry does not require the pore forming toxin Slo. Mouse models have also been utilised in clarifying SlaA as a virulence factor in *S. pyogenes* as comparison between the isogenic mutant strain and the wild type showed that the presence of SlaA increased tissue destruction and systemic spread *in vivo*. Larger lesions and a more rapid mortality were also observed in mice after subcutaneous injection (Sitkiewicz, et al. 2006). A cynomolgus macaque model of pharyngitis has also been employed to investigate the role of SlaA in upper respiratory tract colonisation. SlaA was thus revealed to be a key colonisation factor as the SlaA mutant strain was shown to have severely compromised colonisation ability (Sitkiewicz, et al. 2006). It is as yet unknown whether SlaA contributes to human invasive diseases such as necrotizing fasciitis or if an inflammatory cascade is triggered by SlaA *in vivo*. A summary of putative virulence SlaA effects on eukaryotic cells is shown in figure 1.10

The *S. equi* homologues of SlaA are the subject of Chapter 4

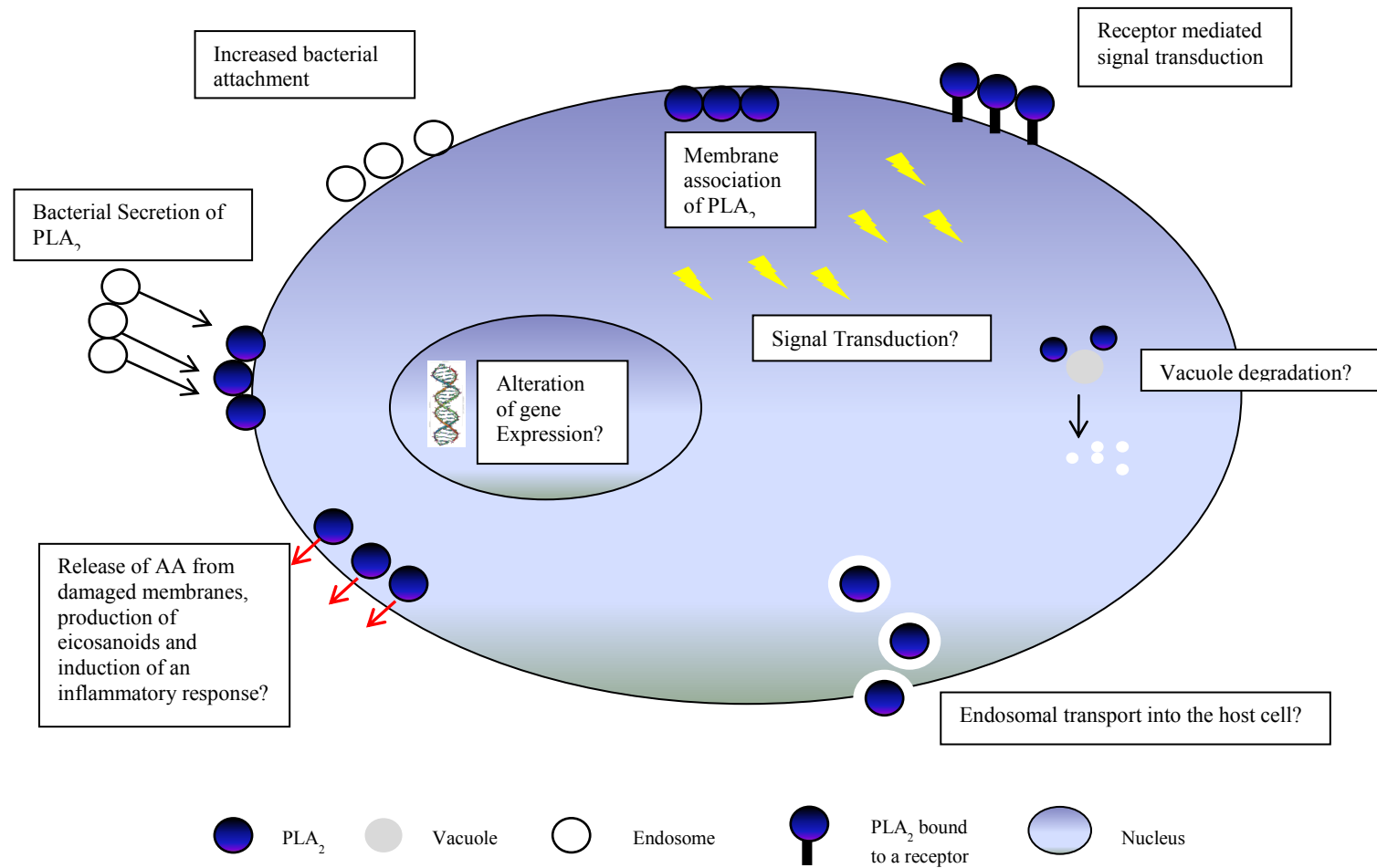


Figure 1.10: The demonstrated and putative effects of SlaA from *S. pyogenes* on the eukaryotic cell.

Figure originally adapted from (Sitkiewicz, et al. 2007).

1.9. Acid phosphatase

The ability to dephosphorylate organic compounds is a key requirement for bacteria, with most bacteria encoding several enzymes (known as phosphohydrolyases or phosphatases) which are able to hydrolyse phosphoesters or phosphoanhydride bonds (Rossolini *et al.* 1998). In many cases, these enzymes can either be membrane bound or secreted outside the plasma membrane in soluble form. Phosphatases were initially classified based on biophysical properties such as pH optimum and substrate profile but have now been grouped following the availability of molecular sequence data and the identification of conserved sequence motifs (Rossolini *et al.* 1998).

Despite being of considerable scientific interest, little information regarding the nature of bacterial phosphatases exists with most information derived from a small set of bacterial models (*E. coli* or *Salmonella enterica* ser. typhimurium). Furthermore the phosphatase pattern of activity is indicated to be highly variable with a large degree of variation shown even in closely related species (Thaller *et al* 1995).

1.9.1. Bacterial non-specific acid phosphatase (NSAP)

A group of phosphatases of particular interest are the bacterial non-specific acid phosphatase (NSAP). The term 'NSAP' initially indicated bacterial enzymes that show optimal activities at acidic to neutral pH conditions and are active towards a variety of different structurally unrelated substrates (Rossolini, et al. 1998).

NSAP activity was first identified in periplasm of *E. coli* (Dvorak 1967). However characterisation of a bacterial NSAP was not available until the studies of PhoN (non-specific acid phosphatase I) and AphA (non-specific phosphatase II), produced by *S. enterica* ser. *Typhimurium* (Kier 1977; Uerkvitz 1998). Subsequent studies indicated similar proteins to AphA and PhoN in a widespread variety of microbial taxa (Rossolini *et al.* 1998).

In contrast some secreted phosphatases have specialization towards specific substrates such as AppA from *E. coli* to 2':3'-cyclic phosphodiesterase (Golovan *et al.* 2000). Several NSAPs from different molecular families can be found encoded by the same organism, which suggests that different classes are responsible for different physiological roles. NSAPs have now been characterised into three molecular families (A, B and C) (Reilly and Calcutt 2004; Rossolini *et al.* 1998; Thaller *et al.* 1998).

1.9.1.1. Class A and B NSAP

Molecular class A and B were proposed following investigations into PhoC (Thaller *et al.* 1994) and NapA (Thaller *et al.* 1995) of *Morganella morganii* respectively. Class A NSAPs share conserved sequence motifs with other eukaryotic phosphatases, with the presence of

conserved residues (G-S-Y-P-S-G-H-[TA]) (Rossolini *et al.* 1998) suggesting an active site. They are also secreted monomeric or oligomeric proteins and consist of a 25-27 kDa polypeptide components. In comparison Class B NSAPs are 100-kDa secreted homotetrameric metallo-proteins with a 25-kDa component. While similar in polypeptide size, class A and B NSAPs are unrelated in terms of sequence homology.

Differentiation of class A and B NSAPs can be achieved via zymographic investigations. Class A acid phosphatases can be detected by zymography i.e renaturing SDS-PAGE using various substrates, usually 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) (Thaller, 1995). BCIP however has not yet been shown to be hydrolysed by any molecular class B NSAP. Resistance to inhibition by EDTA can also be used to differentiate class A to class B (Rossolini *et al.* 1998).

1.9.1.2. Class C NSAP

Of particular interest is the class C NSAP which, are characterized by the presence of four invariant aspartic acid residues within a signature motif (Thaller *et al.* 1998). These are structurally and evolutionary similar to the Class B NSAPs but typically contain an amino-terminal signal sequence type found in putative membrane-bound lipoproteins, hence are not secreted across the cytoplasmic membrane as soluble periplasmic proteins (Rossolini *et al.* 1998).

The first identified class C NSAP was OlpA of the clinically significant *Chryseobacterium* (formerly *Flavobacterium*) *meningosepticum*, which was detected by zymography screening of a group of non-enterobacterial species (Thaller *et al.* 1998).

Following detection and identification of OlpA, two other Class C NSAPs were identified through BLAST searches: e(P4) of *Haemophilus influenzae* and LppC of *Streptococcus equisimilis* (Gase *et al.* 1997; Green *et al.* 1991). Since this initial identification, e(P4) of *H. influenzae* was purified and the first Class C NSAP to be characterised in any detail (Reilly and Smith 1999). This has since been followed by several other NSAPs including LppC (Malke 1998) which was also found to be expressed in *S. pyogenes* (Gase *et al.* 1997). Recently acid phosphatase activity has also been detected in whole cell assays of *S. equi* strain 9682 and *S. zooepidemicus* strain 7023 with both organisms showing western blot cross-reaction with anti-LppC polyclonal antibody, although neither organisms displayed NSAP activity zymographically (Hamilton *et al.* 2000).

1.9.2. Putative NSAP role in virulence

A well established function of NSAP regards the acquisition of nutrients in the form of inorganic phosphate (Pi) and organic by-products that are able to be transported across the membrane (Rossolini *et al.* 1998). However some NSAPs have evolved specialized functions relating to microbial virulence. A common implication of acid phosphatases as a virulence factor lies in aiding bacteria to survive inside phagocytes (Baca *et al.* 1993; Felts *et al.* 2005; Reilly *et al.* 1996; Saha *et al.* 1985).

Coxiella burnetii is a well described pathogen known to proliferate within host acidic phagolysosomes (Akporiaye 1983). Despite possessing superoxide dismutase and catalase which are likely to protect it from the action of host phagocytes (Akporiaye 1983b), it has been shown that *C. burnetii* possesses acid phosphatase activity which is also likely responsible for inhibiting human neutrophils (Baca 1993).

Of particular interest and controversy are the respiratory-burst inhibiting acid phosphatases of *Legionella micdadei* (Das *et al* 1986) and *Francisella tularensis* (Saha *et al* 1985). A strain of *F. tularensis* (ATCC 6223, B38) in particular has been shown to exhibit very large amounts of a tartrate-resistant NSAP activity, AcpA (Reilly, *et al* 1996). AcpA was also shown to be unique in that it was easily released in soluble form from the bacterium. AcpA mutants have also been shown be more susceptible to intracellular killing than wild-type strains in the THP-1 human macrophage-like cell lines. An increased survival was also shown in mice infected with AcpA mutants compared to the wild-type (Mohapatra 2007). However conflicting evidence has been presented to suggest that acid phosphatases do not contribute to the pathogenesis of Type A *F. tularensis* (Child *et al* 2010).

The P4 NSAP has also demonstrated potential as a vaccine target in *H. influenza* as it is widely conserved amongst *H. influenza* strains (Green *et al.* 1991). Reduced intranasal colonization has also been demonstrated in a mouse model with a recombinant form of lipidated P4 protein (Hotomi *et al.* 2005).

The *S. equi* LppC acid phosphatase is the subject of Chapter 5.

2. Materials and methods

2.1. Chemicals

Solution recipes are included in Appendix I or are presented in the text. Standard laboratory chemicals and reagents were obtained from either Sigma-Aldrich[®] (UK) or Fisher Scientific (UK) unless stated otherwise. Unless otherwise stated all solutions and buffers were made using 18.2 MΩ/cm H₂O purified using a Milli-Q plus 185 water purification system, and stored at room temperature.

2.2. Commercial kits

A number of commercial kits were used during this research and a list can be found in Appendix II. Manufacturer's instructions were followed without modification unless otherwise stated.

2.3. Equipment

A list of equipment used in this study can be found in Appendix II.

2.4. Media and solutions

Media recipes are included in Appendix I or are presented in the text. All media were prepared using 18.2 MΩ/cm H₂O. Liquid media was stored at room temperature and solid media at 4°C.

2.5. Bacterial strains, storage and growth conditions

Streptococcus equi 4047 and *Streptococcus zooepidemicus* H70 were from the collection of Prof. Iain Sutcliffe and were originally kindly provided by Dr Neil Chanter (The Animal Trust [Newmarket, Suffolk]).

Glycerol bacterial stocks (0.5 ml overnight culture and 0.5 ml 50% (v/v) glycerol) were aliquoted into Nalgene tubes (Sigma-Aldrich[®], UK) and stored at -80°C. For agar plate storage, *S. equi* and *S. zooepidemicus* streak plates were grown using Todd Hewitt agar (Oxoid, UK) overnight at 37°C in a candle jar and stored at 4°C. *S. equi* and *S. zooepidemicus* broth cultures were grown without agitation at 37°C. To confirm the nature and purity of cultures gram-staining and Lancefield group C testing (Prolab Diagnostic) was undertaken.

2.6. *E. coli* bacterial strains and vectors

All bacterial strains were maintained at -80°C in 50% (v/v) glycerol, and were used as detailed in Table 2.1.

Table 2.1: Bacterial strains and plasmids used in this study

<u>E. coli strain/Vector</u>	<u>Characteristics</u>	<u>Application</u>	<u>Reference/Supplier</u>
One Shot™ TOP10	$F^- mcrA \Delta (mmr-hsdRMS-mcrBC)$ $\Phi80lacZ\Delta M15$ $\Delta lacX74 deoR recA1$ $araD139 \Delta (ara-leu)7697 galU galK$ $rspL endAI nupG$	Cloning host	Invitrogen™
BL21 (DE3)	$F^- ompT hsdS_B (r_B^- m_B^-) gal dcm$	Expression host	Novagen - (Wood 1966)
B834 (DE3)	$F^- ompT hsdS_B (r_B^- m_B^-) gal dcm met$ (DE3)	Expression host	Novagen
C41 (DE3)	$F^- ompT hsdS_B (r_B^- m_B^-) gal dcm lacY1$	Expression host	Lucigen - (Miroux and Walker 1996)

C43 (DE3)	F- <i>ompT hsdS_B</i> (r_B^- m_B^-) <i>gal dcm lacYI</i>	Expression host	Lucigen - (Miroux and Walker 1996)
Tuner™	F- <i>ompT hsdS_B</i> (r_B^- m_B^-) <i>gal dcm lacYI</i>	Expression host	Novagen
BL21 (DE3)pLysS	F- <i>ompT hsdS_B</i> (r_B^- m_B^-) <i>gal dcm met</i> (DE3) pLysS (Cam ^R)	Expression host	Novagen
pCR [®] -Blunt	Kan ^r , Zn ^r , T7 <i>lacZα-ccdB</i>	Cloning vector	Invitrogen™
pET28a	Kan ^r , T7 <i>lac</i> , <i>lacI^q</i>	Expression vector	Novagen

2.7. Antibiotics

All antibiotics were purchased from Sigma-Aldrich[®] (UK) and stock solutions were filter sterilised (Table 2.2).

Table 2.2: Antibiotics used in this study

<u>Antibiotic</u>	<u>Stock Concentration and Solvent</u>	<u>Storage</u>
Chloramphenicol (Melford)	30 mg/ml in 100% Ethanol	-20 °C
Kanamycin (Melford)	10 mg/ml in 18.2 MΩ/cm H ₂ O	4 °C for less than 7 days or -20 °C for long term storage

2.8. Enzymes

All enzymes and reaction co-constituents were maintained at -20 °C, in 50 % (v/v) glycerol unless otherwise stated. A list of general enzymes used in this study is shown in Table 2.3.

Table 2.3: Enzymes and co-constituents

<u>Enzyme (Stock concentration)</u>	<u>Co-constituents</u>	<u>Supplier</u>
KOD DNA polymerase	Amplification buffer (1 x): Patented	Novagen
RNase (10 mg/ml)	Added to Qiagen Resuspension buffer: 10 mM Tris-HCl pH 7.5, 1 mM EDTA, to a final concentration of 100 µg/ml RNase	Qiagen
T ₄ DNA ligase	Ligation buffer (1 x): 50 mM Tris-HCL pH 7.5, 10 mM MgCl ₂	New England Biolabs
Restriction enzymes: <i>Nde</i> I <i>Eco</i> RI <i>Xho</i> I	A list of buffer constituents used for each restriction enzyme is given in Appendix I	New England Biolabs

Acid Phosphatase, from Potato	18.2 MΩ/cm H ₂ O – Stored at -20 °C	Sigma- Aldrich®
PLA ₂ , from Bee	10 mM Tris-HCL, pH 7.5, containing 0.9% KCl and 0.2% BSA – Stored at - 20°C	Caymen Chemical

2.9. Oligonucleotides

Oligonucleotides were synthesised by TAGN, International Centre for Life, Newcastle, UK. Lyophilised oligonucleotides were resuspended in an appropriate volume of sterile 18.2 MΩ/cm H₂O to produce a stock concentration of 100 pmol/μl.

Oligonucleotides were stored at -80°C and working stock concentrations were 10 pmol/μl in 18.2 MΩ/cm H₂O, stored at -20°C. Oligonucleotides used in this study are listed in Table 2.4. The plasmids generated in this study are listed in Table 2.5.

Table 2.4: Nucelotide primers used in this study

<u>Name</u>	<u>Nucleotides (5'-3')</u>	<u>GC content (%)</u>	<u>Tm (°C)</u>	<u>Restriction enzyme</u>
<i>Const_2180_R*</i>	ATC CAT TAA GAA TCA AAT GAA AAG	25.0	52.5	N/a
<i>Const_2190_R*</i>	AGG ACA GCC TGT TTA GCT TGA	47.6	57.9	N/a
<i>Full_2180/90_X_R</i>	CTC GAG TTA AGT TGC TGG TAA TTG TTT TTT AGG	36.4	64.5	XhoI
<i>Full_2180_N_F</i>	CAT ATG GAT ATG AGT GCG GAC GTA AA	42.3	61.6	NdeI
<i>Full2180/2190_E_R</i>	GAA TTC TTA AGT TGC TGG TAA TTG TTT TTT AGG	30.3	62.0	EcoRI
<i>Full2190_N_F</i>	CAT ATG GAT ATG AAA AGC GAT	38.5	60.1	NdeI
<i>Part_2190_E_R</i>	GAA TTC TTA CTC AGC CAT TAA TTG TTT TTC	30.0	59.9	EcoRI
<i>Seq0849f</i>	CAT ATG GAA GGG ATA AAT GAT AAA A	40.0	48.0	NdeI
<i>Seq0849f</i>	CTC GAG AAT AAA TAA TAT TCC CCA AA	52.0	54.0	XhoI
<i>Seq2155f</i>	CAT ATG GCA GAT ACT GCA CCT GCA AGT	56.2	62.4	NdeI
<i>Seq2155r</i>	CTC GAG ATT TCC CAA AAG TCC AGG TAA	56.0	57.0	XhoI
<i>4047LppCf</i>	TGCCATAGGAATTCACGACAAATGGAGAGAAAAAA	37.1	66.0	NdeI

<i>4047LppCr</i>	TGCCATAGCTCGAGTTAATAAGCCTCTAAAGCCTT	42.9	68.3	XhoI
<i>H70LppCf</i>	TGCCATAGGAATTCGCGACAAATGGAGAGAAA	43.8	66.9	NdeI
<i>H70LppCr</i>	TGCCATAGCTCGAGTTAATAAGCCTCTAAAGCT	42.4	67.0	XhoI
<i>H70LppCSeqF</i>	CCTTGTCTTGACTAGGAG	52.4	59.8	N/a
<i>H70LppCSeqR</i>	TGAGAGGTGGATAAATAACCG	42.9	55.9	N/a

Table 2.5 : Recombinant plasmids generated in this study.

<u>Plasmid name</u>	<u>Constituent</u>
pb2190	pCR'Blunt / <i>S. equi</i> , seq2190
pb2190 N	pCR'Blunt / <i>S. equi</i> , seq2190 (truncated gene)
Pb2180	pCR'Blunt / <i>S. equi</i> , seq2180
Pb0849	pCR'Blunt / <i>S. equi</i> , seq0849
Pb2155	pCR'Blunt / <i>S. equi</i> , seq2155
PbSeLppC	pCR'Blunt / <i>S. equi</i> , seq0346
PbSzLppC	pCRBlunt / <i>S. zooepidemicus</i> , szo16870
p2190	pET28a / <i>S. equi</i> , seq2190
p2190 N	pET28a / <i>S. equi</i> , seq2190 (truncated gene)
P2180	pET28a / <i>S. equi</i> , seq2180
P0849	pET28a / <i>S. equi</i> , seq0849
P2155	pET28a / <i>S. equi</i> , seq2155
PSeLppC	pET28a / <i>S. equi</i> , seq0346
PSzLppC	pET28a / <i>S. zooepidemicus</i> , szo16870

2.10. Spectrophotometry

DNA concentration was quantified accurately using an ultraviolet (UV) spectrophotometer (Helios Alpha UV-vis spectrophotometer from Thermo Scientific).

DNA was quantitatively determined by preparing a dilution sample into 50 µl and transferring the dilution to a quartz cuvette. The UV spectrophotometer was previously zeroed with 18.2 MΩ/cm H₂O between the wavelengths of 200 and 300 nm. The optimal wavelength for the absorbance of DNA was assumed to be 260 nm and protein at 280 nm. Double stranded DNA at 50 µg/ml has an A₂₆₀ reading 1.0, allowing the concentration and purity was determined using the below formula.

$$[\text{dsDNA}] \text{ } \mu\text{g}/\mu\text{l} = (A_{260}) \times 50 \times \text{dilution} / 1000.$$

$$\text{DNA purity ratio} = (A_{260}) / (A_{280})$$

A purity ratio of 1.5-1.8 was accepted as minimal protein concentration.

Protein Concentration was quantified accurately by preparing a dilution of 50 µl and transferring the dilution to a quartz cuvette. The UV spectrophotometer was previously zeroed with buffer appropriate to the specific protein at 280 nm. The concentration was determined using the estimated absorbance coefficient of each protein, determined bioinformatically using ExPASy ProtParam tool (<http://expasy.org/tools/protparam.html>).

The formula below was then used when pathlength was 1 cm.

$$\text{Concentration (mg/ml)} = \text{Absorbance at 280 nm} / \text{absorbance coefficient}.$$

2.11. Preparation of chemically competent *E. coli* Top10

50ml of LB media was inoculated with a loop full of *E. coli* TOP10 cells from a glycerol stock and grown at 37°C and 200 rpm until an optical density (O.D) of 0.35-0.4 was observed at 600 nm. 25 ml of this culture was then transferred to a sterile tube and kept on ice for 10 minutes. The culture was then centrifuged at 4000 rpm for 10 minutes at 4°C, the supernatant was removed and the pellet retained. The pellet was re-suspended in 7.5 ml of ice cold MgCl₂-CaCl₂ solution and centrifuged at 4000 rpm for a further 10 minutes at 4°C. The pellet was retained and further re-suspended in 0.5 ml of ice cold 0.1M CaCl₂. The re-suspended pellet was then retained on ice for 2 hours to allow for chemical competence to be induced.

2.12. DNA methods

2.12.1. **Streptococcal genomic DNA extraction**

One loop full of *S. equi* and *S. zooepidemicus* glycerol stock was placed in a 0.5 ml culture of Todd-Hewitt media and grown overnight in a candle jar with no agitation at 37°C. The culture was then centrifuged at 5000 x G (7500 rpm for 10 minutes), the resulting pellet retained and supernatant discarded. To obtain the genomic DNA a DNeasy kit (Qiagen) was used following the protocol supplied by the manufacturer. The volume of DNA obtained was 100 µl.

2.12.2. Restriction enzyme digestion

DNA was digested using restriction enzymes with their appropriate buffers following the manufacturer instructions (New England Biolabs). Each μg of DNA was cut with 5 units of enzyme at 37°C for > 2 hours. Double digests were performed with buffers compatible with both enzymes. Where necessary, Bovine Serum Albumin (BSA) was added at a ratio of BSA to buffer of 1:1. Each digest was ensured to have a glycerol concentration not exceeding 10 % (w/v).

2.12.2.1. PCR

All PCR reactions used *KOD* DNA polymerase. The PCR reaction master mixture's for *KOD* DNA polymerase each contained 10 pmols of each primer, 200 μM of deoxynucleotide triphosphates (dNTPs), MgCl_2 in a range of 1.5-3 mM concentrations, 1 μl of template DNA, 1 unit of *KOD* polymerase and 10x *KOD* reaction buffer. The reaction master mix was made up to a total volume of 50 μl with 18.2 Ω/cm H_2O . Reaction mixtures were prepared in sterile microcentrifuge tubes on ice. Samples were loaded into a temperature gradient thermocycler (Analytik-jena). The general reaction employed for standard PCR was set up as detailed in Table 2.6.

Table 2.6: Reaction conditions for standard PCR.

<u>Temperature (°C)</u>	<u>Duration (min: sec)</u>	<u>Cycles</u>
94.0	02:00	First Denaturation
94.0	00:15	25
Annealing temperature†	00:30	
70.0	00:30*	
70.0	10:00	Final Extension
10.00	Hold	

* Length of extension was determined by expected DNA product size, an extension time of 30 seconds was employed per 500 bp.

† Appropriate annealing temperature for each primer set is detailed in

Table 2.4 and adjusted where nessesary for best results.

2.12.3. Agarose gel electrophoresis of DNA samples

DNA samples were separated by agarose gel electrophoresis. 10 µl samples were mixed with 2 µl of 10x loading buffer (0.1% bromophenol blue, 0.1% xylene cyanol, 60% glycerol, 60 mM EDTA), prior to being loaded onto the agarose gel. The samples were electrophoresed at 100 mA for 20-30 minutes on a 1 % agarose gel in 1x TAE buffer (40 mM Tris-acetate, 10mM EDTA) with one lane containing a 10 kb DNA size standard. The tank buffer was also composed of 1x TAE buffer. Following electrophoresis, the DNA products were stained for 5-10 minutes in 0.5 µg/ml ethidium bromide and visualised using UV light using a gel documentation system (Bio-Rad, UK).

2.12.4. Cloning of DNA fragments into pCR Blunt

Blunt-ended PCR products were produced by amplification with a proofreading DNA polymerase (*KOD* polymerase) and were ligated into pCR Blunt Vector using T₄ DNA ligase. The reaction mixture was prepared according to the manufacturer's instruction (Table 2.6).

Recombinant plasmids were transformed into chemically competent *E. coli* TOP10 cells and were selected for on Luria-Bertani (LB) plates containing kanamycin.

Table 2.7: components of PCR[®]-Blunt ligation reaction.

<u>Component</u>	<u>Volume</u>
PCR [®] -Blunt vector (linearised; 25 ng)	1 µl
10 x Ligation buffer (with ATP)	1 µl
Blunt PCR product	1-5 µl
18.2 MΩ/cm H ₂ O	To total 10 µl
T ₄ DNA Ligase (4U/µl)	1 µl

2.12.5. Plasmid isolation

Plasmid DNA was extracted from the *E. coli* transformant cultures and detection of the insert was performed by undertaking appropriate restriction enzyme digestions. Following identification of the correct sized product, plasmid stocks were purified from the cultures of *E. coli* transformants using a QIAprep spin mini-prep kit (Qiagen, UK) following the protocol supplied by the manufacturer.

For large scale plasmid extractions an alternative STET preparation protocol was employed. A 1.5 ml culture of each *E. coli* strain containing a plasmid was centrifuged at 16,000 rpm for one minute, the supernatant was removed and resuspended in 150 µl of STETS buffer (Appendix I). A total of 100 µg of lysozyme was added to each sample and incubated at room temperature for 15 minutes. The sample was then boiled, centrifuged at 16,000 rpm for 15 minutes and the pellet subsequently removed. The supernatant sample was mixed with 150 µl of chilled isopropanol and stored at -20 °C for around one hour. Following incubation at -20 °C the sample was immediately centrifuged at 16,000 rpm for 5 minutes and a total of 500 µl of ice-cold ethanol (100%) was added. The sample was centrifuged at 16,000 rpm for 3 minutes and the supernatant removed. To obtain the DNA from each sample, the remaining pellet was resuspended in 18.2Ωcm H₂O containing 100 µg of RNase.

2.12.6. Cloning of fragments into pET28a

A total amount of 1500 ng of pET28a vector was cut with appropriate restriction enzymes and concentrated via ethanol precipitation, re-suspended in 10 μ l of 18.2 Ω /cm H₂O. Strains carrying plasmids of pCR Blunt vector with the desired DNA product cloned were inoculated into 5 ml of LB with appropriate antibiotic. Plasmid DNA was obtained using a Qiagen mini-prep kit (see above). pCR Blunt vectors with the desired DNA insert were then sequentially digested with appropriate restriction enzymes (Table 2.4).

Digested DNA fragments were separated by electrophoresis and extracted using the Gel Extraction kit (Quigen). Cut purified plasmid DNA and cut purified pET 28a vector was then ligated using T₄ DNA Ligase. The reaction mix employed is detailed in Table 2.8.

Table 2.8: components of pET vector ligation.

<u>Component</u>	<u>Volume</u>
pET vector cut with appropriate restriction enzyme (100 ng/μl)	1 μl
10 x Ligation buffer	1 μl
Fragment cut from pCR [®] -Blunt vector and gel purified (100 ng/μl)	1 μl
18.2 MΩ/cm H ₂ O	6 μl
T ₄ DNA Ligase (3 U/μl)	1 μl

2.12.7. DNA Sequencing

DNA samples were prepared at a concentration of 30-100 ng/μl in 30 μl aliquots. Sequencing was undertaken by GATC Biotech, London.

2.13. General protein methods

2.13.1. SDS-PAGE

Polyacrylamide gel electrophoresis was carried out using the method of Laemmli (1970). SDS prepared cell extracts or protein purification fractions (10 μl - 20 μl) were boiled for 4 minutes and separated by electrophoresis (using a polyacrylamide gel with a voltage of 200V) until the dye fronts left the gel (approx. 60 min). This was undertaken using a Bio-Rad mini-Protean II system (Bio-Rad, UK). Lanes containing molecular weight standards were included to give an approximate indication of the size of specific proteins. SDS-PAGE gels consisted of a stacking gel (5%) and resolving gel (12 - 17% depending on target protein size), prepared as indicated in Table 2.9 and table 2.10 respectively.

Following electrophoresis, SDS-PAGE gels were stained in coomassie blue dye (see appendix I) for around 20 minutes and placed in de-stain (see appendix I) for around 6 hours or overnight. Visualisation was achieved using a 2D scanner (Bio Rad, UK).

Table 2.9: Composition of SDS-PAGE stacking gel

<u>Component</u>	<u>12% Resolving gel (Volume)</u>	<u>15% Resolving gel (Volume)</u>	<u>17% Resolving gel (Volume)</u>
40% (w/v) solution (37.5:1 acrylamide: bisacrylamide) (Fisher Scientific)	3.0 ml	3.75 ml	4.25 ml
Solution B*	2.5 ml	2.5 ml	2.5 ml
18.2 MΩ/cm H ₂ O	4.5 ml	3.75 ml	3.25 ml
10% (w/v) Ammonium persulphate (APS)(Fisher Scientific)	50.0 µl	50.0 µl	50.0 µl
Tetramethylethylenediamine (TEMED) (Fisher Scientific)	10.0 µl	10 µl	10 µl

* see appendix I

Table 2.10: Composition of SDS-PAGE stacking gel

<u>Component</u>	<u>12% Stacking gel</u> <u>(Volume)</u>
40% (w/v) solution (37.5:1 acrylamide: bisacrylamide) (Fisher)	0.5 ml
Solution C*	1.0 ml
18.2 MΩ/cm H ₂ O	0.5 ml
10% (w/v) APS (Fisher)	30.0 µl
TEMED (Fisher)	10.0 ul

* see appendix I

2.13.2. Protein extraction

Protein as a cell-free extract was obtained from cultures expressing the target protein by centrifugation of culture at 4000 x g. The supernatant was then removed and re-suspended in Start buffer (See appendix I). The re-suspension was then sonicated at ~15 kHz, 14 mA for 2 minutes in 10 second bursts followed by centrifugation at 24,000 x g for 20 minutes at 4°C. This final supernatant was then removed and either purified or analysed by SDS-PAGE.

2.13.3. Concentration and buffer exchange of proteins

Proteins were concentrated by performing centrifugation at 4000 x g at 4 °C using Vivaspinn concentration units with an appropriate volume and molecular weight cut-off (Viva Science, USA). Proteins were also buffer exchanged by dilution within concentrator units.

2.13.4. Protein purification by Immobilized Metal Affinity Chromatography (IMAC)

Purification of recombinant protein containing a His₆-tag was undertaken under native conditions. High yield purification of protein was achieved using an automated FPLC[®] system gradient elution technique. The column was prepared by adding ~ 50 ml Sepharose[™] chelating fast flow resin to a C series column (GE Healthcare, UK) and connected to a FPLC[®] column. 1M NiSO₄ solution was initially passed through the resin until a uniform

colour was achieved. The column was then washed with ~200 ml of Start Buffer (See appendix I) and equilibrated. Appropriate volumes of cell free extract (10-20 ml) was then loaded onto the column at a rate of 4 ml/min and washed through with 100 ml of appropriate buffer at a flow rate of 5 ml/min generated by the FPLC[®] system. Finally, the protein was eluted at a flow rate of 5 ml/min using a linear gradient of extending from 10 mM to 500 mM imidazole with start buffer as diluents (Elution buffer) for 50 minutes.

Monitoring at A₂₈₀ using an inline UV spectrometer allowed the presence of target proteins to be confirmed, and samples were collected in 5 ml volumes on a fraction collector carousel. The purity of protein was confirmed by analysing 20 µl of each 5 ml fraction containing target protein by SDS-PAGE.

2.13.5. Protein purification by Fast Protein Liquid Chromatography (FPLC)

Concentrated protein samples were loaded onto a gel filtration column by injection and eluted over 120 min with gel filtration buffer (See appendix I) running at 1 ml/min. The gel filtration column was of dimensions 1.6 cm (i.d.) x 60 cm (L), with a bed volume of 120 ml and prepared by equilibration at 0.5 ml/min with 120 ml of gel filtration buffer. The point of target protein elution was identified using an inline UV spectrophotometer at A₂₈₀ and samples were collected on a fraction collector in 5 ml volumes. The purity of protein was confirmed by analysing 20 µl of each 5 ml fraction containing target protein by SDS-PAGE.

2.13.6. Western blotting

SDS-PAGE gels for Western blotting were incubated in transfer buffer (25 mM Tris, 192 mM glycine, 10% analytical grade methanol) for 15 minutes, as were the nitrocellulose membranes and blotting paper. Protein transfer onto the nitrocellulose membrane was performed using a Transblot semi-dry system (BioRad, UK). The current applied was 0.6 milliamps per nitrocellulose membrane transfer for 90 minutes. Efficient protein transfer was confirmed by reversible staining with Ponceau S dye (Sigma-Aldrich®, UK). This also allowed the positions of the protein standards to be marked on the nitrocellulose. Blots were blocked overnight with 5% (w/v) skimmed milk in phosphate buffered saline (PBS) and 0.05% (w/v) Tween 80 (PBST) and then incubated for 90 minutes at room temperature with recommended dilutions of the required antibody (Table 2.11). Blots were extensively washed with several changes of PBST and incubated with an appropriate dilution of antisera (Table 2.12). After extensive washing with several changes PBST, blots were developed with Nitro Blue Tetrazolium/5-Bromo-4-Chloro-3-Indolyl Phosphate (NBT-BCIP) (Zymed Laboratories, USA), the alkaline phosphatase substrate.

Table 2.11: Primary antibodies used in this study

<u>Antibody</u>	<u>Appropriate dilution</u>	<u>Protein detected</u>	<u>Supplied by</u>
Anti-His ₆ -Tag	1/3000	Specific His ₆ -Tagged protein	Sigma-Aldrich®
Anti-LppC	1/1000	LppC acid phosphatase	Dr Horst Malke (Jena University, Germany)
Equine Convalescent serum (Pre infection)	1/1000	Non-Immunoreactive <i>S. equi</i> proteins (Control)	Animal Health Trust, Newmarket, UK
Equine Convalescent serum (Post infection)	1/1000	Immunoreactive <i>S. equi</i> proteins	Animal Health Trust, Newmarket, UK

Table 2.12: Secondary antibodies used in this study

<u>Secondary antibody</u>	<u>Appropriate dilution</u>	<u>Relevant primary antibody</u>	<u>Supplied by</u>
Alkaline phosphate-conjugated goat anti-rabbit IgG	1/30,000	Anti-LppC	Sigma-Aldrich®
Alkaline phosphate conjugated horse IgG	1/10,000	Anti-His ₆ -Tag	Sigma-Aldrich®
Alkaline phosphate conjugated mouse IgG	1/10,000	Equine Convalescent serum (pre and post infection)	Sigma-Aldrich®

2.13.7. Protein crystallisation

Cover slips were initially prepared using liquid Aquasil. Prior to use, cover slips were wiped with a clean cloth. Crystal tray wells were filled with the appropriate buffer solutions from various screens and the coverslips containing drops of protein at a working concentration of 15 mg/ml in three different ratios of 1:1, 1:2 and 2:1 respectively and were inverted onto the surface of the wells which were previously greased. Once the tray of 24 wells was set up, they were incubated at 21.0°C in a crystallisation incubator and monitored for the presence of crystals every week by microscopy. The various screens used and their buffer compositions are included in Appendix III.

2.14. Enzyme assay methods

2.14.1. General kinetic analysis considerations

All kinetic assays were performed using the same set of calibrated pipettes (Gilson[®]) and all enzyme assays were performed in triplicate with mean values presented on the graphs with error bars to represent the range for each data set where necessary. Each assay dataset was also produced using the same aliquot of purified enzyme.

2.14.2. sPLA₂ assay methods

2.14.2.1. SIGMA assay

Initially putative secreted phospholipase activity was investigated using a method supplied by Sigma-Aldrich®

(http://www.sigmaaldrich.com/etc/medialib/docs/Sigma/General_Information/phospholipase_a2_colorimetric.Par.0001.File.dat/phospholipase_a2_colorimetric.pdf).

70 µl of phospholipase assay buffer (500 mM Tris HCL, 10 mM CaCl₂, pH 8.5) was added to 70 µl of 2% (w/v) L-a-phosphatidylcholine (Sigma-Aldrich®) and 1.5% (w/v) Deoxycholate solution. This mixture was then equilibrated to 37°C and 0.5 µl of purified enzyme added. Following addition, the sample was mixed by inversion and incubated at 37°C for a further 5 minutes and 20 µl removed. To the removed 20 µl sample; 150 µl of ether, 20 µl of 2 M hydroxylamine solution and 20 µl of 14 % (w/v) sodium hydroxide solution were added. This was then incubated at 25°C for 20 minutes and finally, 30 µl of 3 N hydrochloric acid solution and 30 µl of 10% (w/v) ferric chloride solution was added. The reaction mixture was mixed and absorbance recorded at A_{570nm}.

2.14.2.2. Price colourmetric assay

A further biochemical assay for putative phospholipase activity was investigated using a colorimetric assay developed by Price (2007). The degradation of a common phospholipase A₂ substrate (phosphatidylcholine) allows for maximal detection of many PLA₂ at physiological pH using an indicator dye. Samples were diluted in a cold saline solution (2 mM HEPES at pH 7.5. 20 µl of each sample was loaded into wells of a round bottom 96 well microplate. 180 µl of an assay mixture was loaded to each assayed sample after pre-warming in a water bath at 37°C. The assay mixture contained 5 mM Triton X-100, 5 mM phosphatidylcholine, 2 mM HEPES, 10 mM calcium chloride, and 0.124% bromothymol blue dye in 18.2 MΩ/cm H₂O at pH 7.5.

Phosphatidylcholine had been dissolved at 160 mg/mL in molecular grade methanol and stored at -70°C prior to use. Stock phosphatidylcholine samples were re-suspended at 45°C immediately before use. The final assay mixture concentration of methanol was 5%.

After the reaction was initiated, the multiwell assay plate was immediately analyzed at 595 nM with an EL 808 Ultra microplate reader (Biotek Instrument Inc). The plate was read at 1 minute intervals at 37°C and rates of change were blank adjusted.

2.14.2.3. Chromogenic substrate release assay

Putative phospholipase activity was investigated using sPLA₂ Assay Kit (Caymen Chemical Company, USA). The assay uses 1, 2-dithio analog of diheptanoyl phosphatidylcholine

(DHT-PC) as a standard substrate and detects free thiols using upon hydrolysis of the thio-ester bond at the *sn*-2 position by PLA₂ using the chromogen DTNB (5, 5'-dithio-*bis*-(2-nitrobenzoic acid)) (Figure 2.1). Assay was performed as instructed by the manufacturer. Prior to assay all protein samples were concentrated in sPLA₂ assay buffer (25 mM Tris-HCL, pH 7.5, containing 10 mM CaCl₂, 100 mM KCl and 0.3 mM Triton X-100). DTNB was reconstituted in 1.0 ml of 18.2Ωcm H₂O per supplied vial to yield 10 mM DTNB in 0.4 M tris-HCl, pH 8.0. The reagent was then stored on ice in the dark and used within eight hours. An alternative substrate, 2-deoxy-2-thio- Arachidonoyl phosphatidylcholine (AT-PC) (see appendix I), was also tested utilizing the same method as above.

All substrates were reconstituted with supplied assay buffer and stored at -20 °C. For investigations regarding Ca²⁺ ion requirements, the assay buffer was modified accordingly.

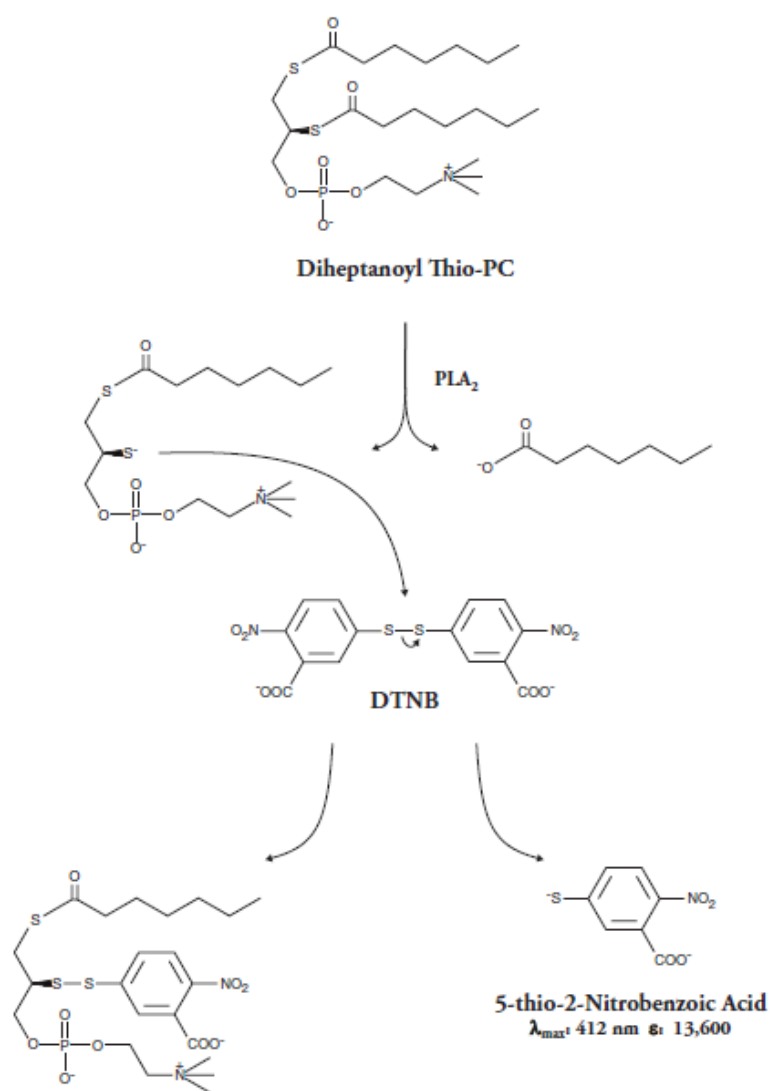


Figure 2.1: Schematic representation of sPLA₂ activity reaction.

Figure obtained from sPLA₂ assay kit booklet, Caymen Chemical Company.

2.14.3. Acid phosphatase assay

Putative acid phosphatase activity was investigated using a method developed by Reilly et al (2009). The assay utilises *para*-nitrophenyl phosphate (*p*NPP) as a chromogenic substrate. The acid phosphatase dephosphorylates *p*NPP and under alkaline conditions the phenolic OH-group is deprotonated resulting in *p*-nitrophenolate (Figure 2.2). A standard 190 µl assay mixture (0.2 M Sodium acetate, pH 5.0, 0.1 mM CuSO₄, 10.0 mM *p*-nitrophenylphosphate (*p*NPP)) was added to 10 µl of sample. The reaction was incubated at 37°C for 10 minutes and stopped by the addition of 100 µl of 0.5 mM NaOH, pH 10.0. The concentration of *p*-nitrophenol produced was measured using an EL 808 Ultra microplate reader (Biotek Instrument Inc, UK) microtiter plate reader at 410 nm. Prior to assay all protein samples were concentrated in buffer containing 0.2 M sodium acetate, pH 5.0. For investigations regarding divalent ion requirement, the standard assay mixture was modified accordingly.

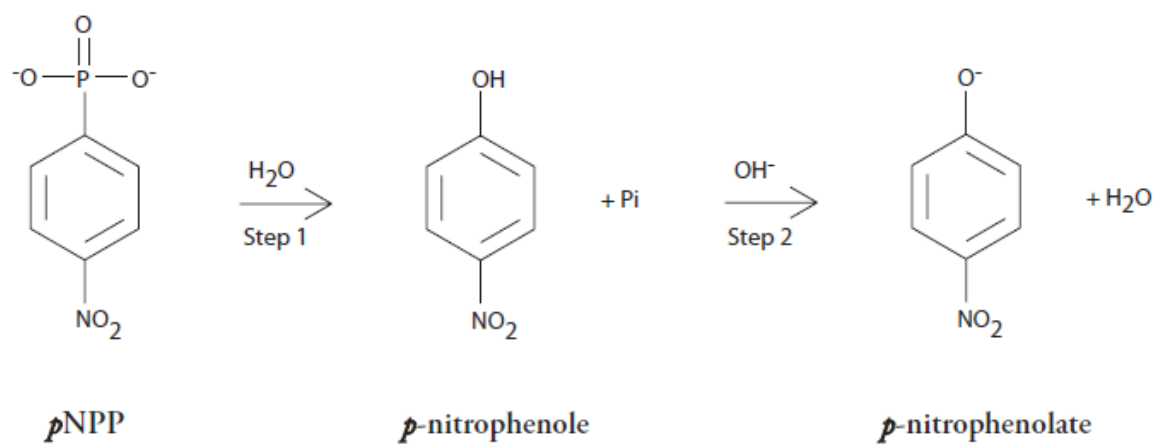


Figure 2.2: Schematic representation of acid phosphatase activity reaction.

Figure obtained from Caymen Scientific.

2.14.4. Phosphate assay

In order to investigate further investigate acid phosphatase activity using the non-chromogenic substrates Uridine 5' monophosphate (5'UMP) and adenosine 5' monophosphate (5' AMP) the phosphate assay developed by Chen (1956) was used to determine the released phosphate.

Assay was performed as above (Section 2.14.3) and assay mixture removed. The total assay volume (300 μ l) of sample was added to 300 μ l of colour reagent (3 M Sulphuric acid, 2.5% w/v ammonium molybdate, 114 mM Ascorbic Acid). Sample and colour reagent was incubated at 37°C for 2 hours and absorbance value was calculated via spectrophotometer at A_{820} . A set of standards was produced using dilutions of 33 mM KH_2PO_4 .

2.15. Calculations of K_m and specific activity

2.15.1. **sPLA₂ (Chapter 4)**

Considering the calculations of sPLA₂, rate was determined by plotting the absorbance values as a function of time to obtain the slope (rate) of the linear portion of the curve. Two points were chosen on the linear portion of the curve and the change in absorbance calculated by the following equation:

$$\Delta A_{414} = \frac{A_{414}(\text{Time 2}) - A_{414}(\text{Time 1})}{\text{Time 2 (min.)} - \text{Time 1 (min.)}}$$

The rate of $\Delta A_{414}/\text{min}$ for the non-enzymatic controls (blank) was determined and subtracted from this value. The following formula was then used to calculate sPLA₂ rate.

$$\begin{aligned} \text{sPLA}_2 \text{ Activity} = & \frac{\Delta A_{414}/\text{min}}{\text{DTNB adjusted extinction coefficient (10.66 mM}^{-1})} \times \frac{0.225 \text{ ml}}{0.01 \text{ ml} \times \text{Sample Dilution}} \\ & = \text{mM of substrate converted per minute} \end{aligned}$$

The values of K_m obtained in this study were calculated by fitting the data directly to the Michaelis-Menten equation using Graphpad (Graphpad software, Inc). Data was displayed as Lineweaver Burk plots (Lineweaver, 1934) and all K_m values were presented in mM.

Regarding specific activity, the obtained values were converted into μmol of substrate converted per min and further converted into $\mu\text{mol}/\text{min}/\text{mg}$ in relation to the amount of enzyme added in each reaction.

2.15.2. Acid phosphatase (Chapter 5)

Considering the calculations of acid phosphatase activity, the value for K_m was obtained as above (section 2.15.3) with the exception of a *p*NPP adjusted coefficient of 9.248 mM^{-1} and adjusted volumes specific to the assay.

For AP assays using non-chromogenic substrates a line of best fit was instead obtained using dilutions of $33 \text{ mM KH}_2\text{PO}_4$ and individual activity values were determined from the standard curve.

2.16. Zymography

Acid phosphatase enzymes can be assayed in gel ('zymography') following electrophoresis (Thaller *et al* 1995). Putative acid phosphatase enzymes were electrophoresed through 12% polyacrylamide gels. Following SDS-PAGE, separated protein was renaturated for 4 hours by incubation in several changes of renaturation buffer ($100 \text{ mM Tris-HCL pH } 7.0$, containing 2 mM MgSO_4 , 0.05 mM ZnCl_2 and $2\% \text{ (v/v) Triton X-100}$). After renaturation the gels were further incubated for 1 hour in several changes of equilibration buffer ($100 \text{ mM sodium acetate, pH } 5.0$).

Activity of acid phosphatases was detected following overnight incubation, at 37°C , in fresh equilibration buffer containing phosphatase substrate, 0.25 mM BCIP , with the addition of 0.25 mM NBT as described by Hamilton *et al.* (2000). The in gel appearance of blue bands indicated the presence of phosphatase activity.

2.17. Bioinformatic analysis

Nucleotide sequences determined in this research were translated into amino acid sequences using the ExPASy translate tool (<http://expasy.org/tools/dna.html>). All 6 frames are read highlighting start and stop codons. DNA and protein sequences were compared to those in the public domain (notably genbank) using the Basic Local Alignment Search tool (BLAST; Altschul *et al* 1990) (<http://blast.ncbi.nlm.nih.gov/>).

The Sanger Centre *S. equi* finished genome BLAST server (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_equi) was used to search for *S. equi* sequences. The Sanger Centre *S. zooepidemicus* finished genome BLAST server (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_zooepidemicus) was used to search for *S. zooepidemicus* sequences. Homologous proteins were aligned and overall percentage identity and similarity calculated using the ClustalW (Larkin *et al* 2007) alignment tool (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). The membrane spanning domains of proteins were predicted using the topology program TMPRED (Hofmann and Stoffel 1993) (http://www.ch.embnet.org/software/TMPRED_form.html) and the putative signal peptide regions were identified using SignalP (Nielsen *et al* 1997) (http://www.cbs.dtu.dk/services/SignalP/signal_P/). Putative lipoprotein signal peptide regions were identified using PRED-LIPO (Bagos *et al* 2008) (<http://biophysics.biol.uoa.gr/PRED-LIPO/>).

Restriction enzyme mapping was performed using Webcutter 2.0 (Maarek *et al* 1997) (<http://bio.lundberg.gu.se/cutter2/>). The program was used to identify restriction enzyme sites

within target sequences so that restriction enzyme sites, which were not present within the sequences to be amplified, could be incorporated into the primers designed.

Putative protein structure prediction was undertaken using phyre² (Kelley 2009) (<http://www.sbg.bio.ic.ac.uk/~phyre/>)

3. Investigation into the cell wall anchored protein SEQ2190 of *S. equi*.

3.1. Introduction

As previously described (section 1.5), proteins identified as cell surface associated via sortase mediated attachment to the cell wall have been previously shown to function as virulence factors with prominent interactions between bacteria and host cells.

As stated in section 1.2, *S. equi* and *S. zooepidemicus* share a very large proportion of conserved proteins but exhibit very different pathogenicity. Expressed proteins unique to each species are therefore of considerable interest in terms of investigating the molecular basis of host specificity with regard to virulence. The putative cell wall anchored protein SEQ2190 protein represents one such unique expressed protein and further bioinformatic analysis (section 3.2) has revealed the N-terminal region of SEQ2190 (1-252 nucleotides of full length protein) to be only present in known strains of *S. equi* and not present in any known strain of *S. zooepidemicus*. Since bacterial evolution typically favours small genomes to reduce the metabolic cost, it is indicated that the retention and conservation of *seq2190* N-terminal in *S. equi* must provide a significant bacterial survival advantage and hence represent a putative pathogenicity characteristic unique to *S. equi*.

Importantly, SEQ2190 has been previously shown to be immunogenic early in the onset of a *S. equi* infection and is conserved amongst a wide variety of *S. equi* isolates (Waller, Mitchell & Robinson – unpublished). Furthermore an 235 amino acid immunoreactive protein fragment has also been identified (Se68.9) (Timoney *et al.* 2007) which is identical to a portion of the SEQ2190 protein (amino acids 376-610).

Aims of the present study

The aims of the present study were:

- a) To produce cloning constructs expressing the SEQ2190 and its paralogue SEQ2180 to allow further investigation of protein structure and function.
- b) To obtain x-ray crystallography structures to allow for structural comparisons to reveal possible structural orthologues.
- c) To produce gene ‘knock-out’ mutants of SEQ2190 and related homologues.

The overall aim is to assess the potential of SEQ2190 as a virulence factor of *S. equi* and a future target for novel therapeutic or preventative strategies.

3.2. Bioinformatic analysis of SEQ2190

The amino acid sequence of the putative surface associated protein SEQ2190 was obtained from UniProt in a FASTA format. The sequence encoding SEQ2190 was analysed for identification of a signal peptide cleavage site by using the SignalP Hidden Markov Models and Neural Networks outputs, with a cleavage site predicted between position 45 and 46. SEQ2190 was annotated as a putative surface attached protein due to the presence of an LPXTG sortase recognition motif within the C-terminal of the mature protein (position 589-593) with alanine present as the X within the motif. A BLAST search of the translated *S. equi* genome database for homologues to SEQ2190 revealed an amino acid sequence with 80% identity, SEQ2180 (Figure 3.1). The full amino acid sequence of SEQ2180 is shorter than SEQ2190 (626 amino acids in length compared to 673) suggesting a truncation of the amino acid sequence or an insertion in SEQ2190. SEQ2180 also differs from SEQ2190 in that the signal peptide cleavage site was instead predicted between position 37 and 38. Notably, sequence differences between the two mature proteins are restricted to the N-terminal region of the two proteins (from amino acid position 3 to position 197 of the mature SEQ2190 protein, Figure 3.1).

A tBLASTn search was undertaken for both SEQ2190 and SEQ2180 against the *S. equi* sub sp. *zooepidemicus* H70 genome project (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_zooepidemicus). Notably, an orthologue of SEQ2180 was found in *S. zooepidemicus* H70. The orthologue, annotated as SZO18890, was found to have 95.0% identity to the full SEQ2180 amino acid sequence (Figure 3.2). In contrast SZO18890 shared only 80% homology with SEQ2190.

A further, shorter (372 amino acids in length) homologue, SZO18870, was also identified and annotated as a pseudogene due to the absence of the signal peptide region or a transcription initiation sequence. The translated pseudogene SZO18870 was shown to have 83.0% and 82.0% identity to the corresponding regions of SEQ2190 and SEQ2180 respectively (Figure 3.3). Database searches for other orthologues to the amino acid sequence of the mature SEQ2190 protein using BLASTP and tBLASTn revealed low but significant homologies to the Fc Gamma receptor protein (M family protein) and M-protein of *S. pyogenes* (Table 3.1). Significantly no homologues to the unique N-terminal regions of SEQ2190 or SEQ2180 were identified. The amino acid sequence of the mature SEQ2190 protein was also analysed for a predicted secondary structure. This was undertaken using Phyre, providing a consensus predicted secondary structure based on the predicted results from Psipred, Jnet and Sspro (Figure 3.4). The amino acid sequence of the mature SEQ2180 protein was also analysed for a predicted secondary structure and as expected demonstrated an almost identical predicted structure (data not shown).

The overall organisation of the SEQ2180 and SEQ2190 locus between the genomes of *S. equi* and *S. zooepidemicus* is shown below (Figure 3.5). Cumulatively these analyses confirm that SEQ2190 is a protein unique to *S. equi* whereas SEQ2180 has a *S. zooepidemicus* orthologue.

SEQ2190	DMKSDAKKVLATLQKGNHKQELKKVLRITGD-DSLLLLWLLVSGDSNLDENKSFYDLSI	59
SEQ2180	DMSADVKKIEQVLQ--SDSGKLKTVLGQISDPNTLFALFAIISADVNI DPESFSF----I	54
	.:*.: .** ... :**. ** *. * :*: *: :*: * *: * .. ** *	
SEQ2190	WNTGSRVGKVISPTHEYFKQESKNTENRQAVYSEFKQERVEERSKKAKEATEALNQKAQLEA	119
SEQ2180	LN-----GPITPLLSYFNPNPNKN-DNRKAVVNEIKKRVS-----ALEQAQKLETKAKIEK	103
	* * : . . **: : . ** :*: ** . *:*: **. * * : : *: **: : *	
SEQ2190	TVKNINEELEKTREGFKVVSENSNKLEKQLMAEKIKTRTAEETAKQAKTDKERAEAEAKK	179
SEQ2180	QIAES--ELEKTRNAFKVTIENLHKLEGLLDTEKQKTRKVEEDYQQAKTDKEKAEAD---	158
	: : *****: . ***. ** :*** * :** ***. . ** :*****:***:	
SEQ2190	AKEEAKTAEGKVKQAETEKRNAEAKARTAE EEAQATADKEKAETEA KKAKEEAKTAKEA	239
SEQ2180	-----KRNAETKARTAE EEAQATADKEKAETEA KKAKEEAKTAKEA	200
	*****:*****	
SEQ2190	AHQEQEKAKQLEQANQQANQRANLAEKSKKDLETQKEKLEQEIKEATEAKNKAEQKLKDL	299
SEQ2180	AHQEQEKAKQLEQANQQANQRANLAEKSKKDLETQKEKLEQEIKEATEAKNKAEQKLKDL	260

SEQ2190	QDSASQGSELSKQLLKEKEELTTKLQELQKQAE EKTTEIEK LKQELEANKQNSGQLGQQE	359
SEQ2180	QDSASQGSELSKQLLKEKEELTTKLQELQKQAE EKTTEIEK LKQELEANKQNSGQLGQQE	320

SEQ2190	QKLQEQLNKVQKELKQKEMELKQAQEQ LKQEQKPHEGGGSDASKARITELEKQVQTLTK	419
SEQ2180	QKLQEQLNKVQKELKQKEMELKQAQEQ LKQEQKPHEGGGSDASKARITELEKQVQTLTK	380

SEQ2190	EKADLSSTLESTKAQLSETQARLSEAQKQLTAAQEKLTTLEAEKTALQHQVETISKQLSE	479
SEQ2180	EKADLSSTLESTKAQLSETQARLSEAQKQLTAAQEKLTTLEAEKTALQHQVETISKQLSE	440

SEQ2190	TRDLSEKEKAALQEQINKLKAEIEQKTKEIEALKQGMQSHQGQEKPKDPKTPETPKDPKT	539
SEQ2180	TRDLSEKEKAALQEQINKLKAEIEQKTKEIEALKQGMQSHQGQEKPKDPKTPETPKDPKT	500

SEQ2190	PEKNDQPQAPEKRSVPWTALTPAKPIDTTKAPKSSAPSPQTGAATPKKQLPATGDTATPS	599
SEQ2180	PEKNDQPQAPEKRSVPWTALTPAKPIDTTKAPKSSAPSPQTGAATPKKQLPATGDTATPS	560

SEQ2190	FFTAAAMAVLASVGVNLSPRRKKNQNNR	628
SEQ2180	FFTAAAMAVLASACALTLSRRKKNQNNR	589
	***** . . *	

Figure 3.1: Alignment of mature SEQ2190 and SEQ2180.

Alignment was performed using ClustalW2.* indicates identical amino acids, : indicates similar amino acids. The LPXTG cell wall anchoring motif in each sequence is underlined.

SEQ2180	MKRQSNIRRLCHQALAATLITGSLIGTGLLGKNKVYV G DMSADVKKIEQVLQSDSGKLKTV	60
SZO18890	MKRQSNIRRLCHQALAATLITGSLIGTGLLGKNKVYV G DMSADVKKIEQVLQSDSGKLKTV	60
	*****★*****	
SEQ2180	LGQISDPNTLFAIFAIISADVNIDPESFSFILNGPITPLLSYFNNPNKNDNRKAVVNEIK	120
SZO18890	LGQISDPNTLFAIFAIISADVNIDPESFSFILNGPITPLLSYFNNPNKNDNRKAVVNEIK	120

SEQ2180	KRVSALEQAQKLETKAKIEKQIAESELEKTRNAFKVTIENLHKLEGLLDTEKQKTRKVEE	180
SZO18890	KRVSALEQAQKLETKAKIEKQIAESELEKTRNAFKVTIENLHKLEGLLDTEKQKTRKVEE	180
	*****:*****	
SEQ2180	DYQQAKT DKEKAEADKRNAETKARTAE EEAQATADKEKAETEAKKAKEEAKTAKEAAHQ	240
SZO18890	DYQQAKT DKEKAEADKRNAETKARTAE EEAQATADKEKAETEAKKAKEEAKTAKEAAHQ	240

SEQ2180	EQEKAKQLEQANQQANQRANLAEKSKKDLETQKEKLEQEIKEATEAKNKAEQKLKDLQDS	300
SZO18890	EQEKAKQLEQANQQANQRANLAEKAKTALETQKQKLEQEIKEATEAKNKAEQKLKDLQDS	300
	*****:*.*****:*****	
SEQ2180	ASQGSSELSKQLLKEKEELTTKLQELQKQAE EKTTEIEKLKQELEANKQNSGQLGQQEQKL	360
SZO18890	ASQGSSELSKQLLKEKEELTTKLQDLQKQAE EKTTEIEKLKQELEANKQNSGQLGQQEQKL	360
	*****:*****	
SEQ2180	QEQLNKVQKELKQKEMELKQAQEQLKQE QKPHEGGGSDASKARITELEKQVQTLTKEKA	420
SZO18890	KEQLNKVQEELKQKEMELEKAKEQLKQE QKPHEG---DDTSKAKIAELEQQVQTLTKEKE	417
	:*****:*****::*:*****.*:***:*:***:*****	

```

SEQ2180      DLSSTLESTKAQLSETQARLSEAQKQLTAAQEKLTTLEAEKTALQHQVETISKQLSETRD 480
SZO18890     NLSSTLESTNAQLSETQAKLSEAQKQLTAAQEKLTTLEAEKVALQHQVETISKQLSEARD 477
              :*****:*****:*****.*****:
              :*****:*****:*****.*****:

SEQ2180      LSEKEKAALQEQINKLKAEIEQKTKEIEALKQGMQSHQGQEKPKDPKTPET-----PKD 534
SZO18890     LSEKEKAALQEQINKLKAEIEQKTKEIEALKQGMQSHQGQEKPKDPKTPETPKAPKVPKD 537
              *****
              *****

SEQ2180      PKTPEKNDQPQAPEKRSVPWTALTPAKPIDTTKAPKSSAPSPQTGAATPKKQLPATGDTA 594
SZO18890     PKTPEKNDQPQAPEKRSVPWTALTPAKPIDTTKAPKSSAPSPQTGAATPKKQLPATGDTA 597
              *****
              *****

SEQ2180      TPSFFTAAAMAVLASACALTLSPRRKKNNR 626
SZO18890     TPSFFTAAAMAVLASVGVLSLSPRRQKNQNNR 629
              *****. .*:*****:*****

```

Figure 3.2: Alignment of SEQ2180 and the orthologue SZO18890.

Alignment was performed using ClustalW2. * indicates identical amino acids, : indicates similar amino acids. The point of predicted signal peptide cleavage is indicated in bold and the LPXTG cell wall anchoring motif in each sequence is underlined.

SEQ2190	MITKKQRKLKRHTTIRRLCHQALAATLITGSLMGTGLLGNKV VYGDMKSDAKKVLATLQK	60
SEQ2180	-----MKRQSNIRRLCHQALAATLITGSLIGTGLLGNKV VYGDM SADVKKIEQVLQ-	51
SZO18970	-----	
SEQ2190	GNHKQELKKVLR TIGD-DSL LLLWLLVSGDSNLDES NKSFDYLSIWNTGSRVGKVIS PTE	119
SEQ2180	-SDSGKLKTVLGQISDPNTL FALFAIISADVNI DPESFSF----ILN-----GPITPLLS	101
SZO18970	-----	
SEQ2190	YFKQESKNTENRQAVYSEFKQ RVEERSKKAKEATEALNQKAQLEATVKNINEELEKTREG	179
SEQ2180	YFNNPNKN-DNRKAVVNEI KKRVS-----ALEQAQKLETKAKIEKQIAES--ELEKTRNA	153
SZO18970	-----	
SEQ2190	FKVVSSENSNKLEKQLMAEKIKTRTAEETAKQAKTDKERAEAEAKKAKEEEAKTAEGKVKQA	239
SEQ2180	FKVTIENLHKLEGLLDTEKQKTRKVEEDYQQAKTDKEKAEAD-----	195
SZO18970	-----	
SEQ2190	ETEKRNAEAKARTAE EEEAKQATADKEKAETEAKKAKEEEAKTAKEAAHQEQEKAKQLEQAN	299
SEQ2180	---KRNAETKARTAE EEEAKQATADKEKAETEAKKAKEEEAKTAKEAAHQEQEKAKQLEQAN	252
SZO18970	-----PK	2
	. :	

SEQ2190	QQANQRANLAEKSKKDLETQKEKLEQEIKEATEAKNKAEQKLKDLQDSASQGSELSKQLL	359
SEQ2180	QQANQRANLAEKSKKDLETQKEKLEQEIKEATEAKNKAEQKLKDLQDSASQGSELSKQLL	312
SZO18970	NRESKKATQAEQAKKALETQKEKLEKDIEQAKEAKNKAEQKLKDLQDSASQGSELSKQLL	62
	:: .::*. **:::* *****:::*. *****	
SEQ2190	KEKEELTTKLQELQKQAEKTTIEIKLKQELEANKQNSGQLGQQEQKLQEQLNKVQKELK	419
SEQ2180	KEKEELTTKLQELQKQAEKTTIEIKLKQELEANKQNSGQLGQQEQKLQEQLNKVQKELK	372
SZO18970	KEKQELTTKLQDLQKQANEKTTIEIKLKQELEANKQSNGQLNQKEQELKNQLDKIQEELK	122
	:**:*****:*****.***.**:*:*:*:*:*:***	
SEQ2190	QKEMELKQAQEQLKQEOKPHEGGGSDASKARITELEKQVQTLTKEKADLSSTLESTKAQ	479
SEQ2180	QKEMELKQAQEQLKQEOKPHEGGGSDASKARITELEKQVQTLTKEKADLSSTLESTKAQ	432
SZO18970	QKEMELEKAKEQLKQEOKPHEY---DDTSKTRITELEQQVQTLTKEKENLSSTLESTKAQ	179
	*****::*:*****.**:*:*****:*****:*****	
SEQ2190	LSETQARLSEAQKQLTAAQEKLTTLEAEKTALQHQVETISKQLSETRDLSEKEKAALQEQ	539
SEQ2180	LSETQARLSEAQKQLTAAQEKLTTLEAEKTALQHQVETISKQLSETRDLSEKEKAALQEQ	492
SZO18970	LTETQAKLSEAQKQLTAAQEKLTTLEAEKVALQQQVDTISKQLSEARDLSEKEKAALQEQ	239
	*:*****:*****.***:*:*****:*****	
SEQ2190	INKLKAEIEQKTKEIEALKQGMQSHQGQEKPKDPKTPETPK---DPKTPEKNDQPQAPEK	596
SEQ2180	INKLKAEIEQKTKEIEALKQGMQSHQGQEKPKDPKTPETPK---DPKTPEKNDQPQAPEK	549
SZO18970	INKLKAEIEQKTKEIEALKQGMQSHQGQEKPKPEPKMPEAPMKPEEPKTPEKKDQPQTPEK	299
	*****:*** **:*: *****:*****:***	

```

SEQ2190      RSVPTALTTPAKPIDTTKAPKSSAPSPQTGAATPKKQLPATGDTATPSFFTAAMAVLAS 656
SEQ2180      RSVPTALTTPAKPIDTTKAPKSSAPSPQTGAATPKKQLPATGDTATPSFFTAAMAVLAS 609
SZO18970     RSIPWTALTTPAKPIDTTKAPKSSAPSSQTGAATPKKQLPATGDTATSSFFTTAAMAVLAS 359
              **.:*****.*****.***:*****

SEQ2190      VGVNLSPRRKKNQNNR 673
SEQ2180      ACALTSPRRKKNQNNR 626
SZO18970     VGVLTSPRRKKD---- 372
              . .*.*****:

```

Figure 3.3: Multiple alignment of SEQ2190, SEQ2180 and the orthologue SZO18970.

Multiple alignments were performed using ClustalW2. * indicates identical amino acids, : indicates similar amino acids.

Table 3.1: Significant orthologues of the mature SEQ2190 protein.

Blast search performed by tBLASTn.

<u>Accession</u>	<u>Annotation</u>	<u>Organism</u>	<u>Identities</u>	<u>Positives</u>	<u>Gaps</u>	<u>E number</u>
gb ACI61926.1	Fc-gamma receptor	Streptococcus pyogenes strain NZ131	100/136 (32%)	147/316 (47%)	54/316 (17%)	2.00E- 11
emb CAA56449.1	M3 Protein	Streptococcus pyogenes	116/385 (30%)	186/385 (48%)	63/385 (16%)	4.00E- 11
NP_665531.1	antiphagocytic M protein, type 3	Streptococcus pyogenes MGAS315	116/385 (30%)	186/385 (48%)	63/385 (16%)	5.00E- 11
gb CAA44062.1	M protein type 1	Streptococcus pyogenes	92/325 (28%)	161/325 (50%)	68/325 (14%)	9.00E- 11

Index

Query Sequence

psipred

jnet

sspro

Consensus

Cons_prob

100. 110. 120. 130. 140. 150. 160. 170. 180.

190. 200. 210. 220. 230. 240. 250. 260. 270.

280. 290. 300. 310. 320. 330. 340. 350. 360.

370. 380. 390. 400. 410. 420. 430. 440. 450. 460.

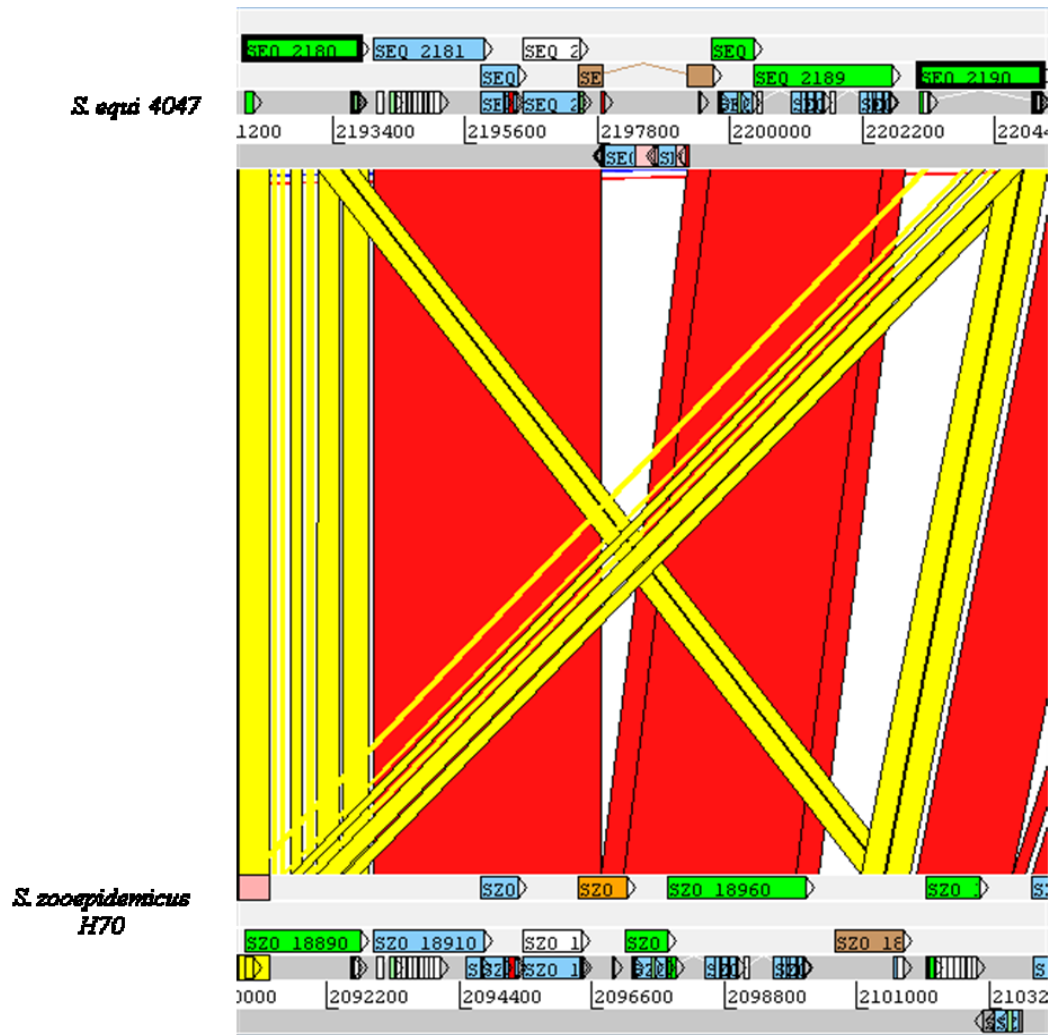


Figure 3.5: Artemis screen capture of the *seq2180* and *seq2190* locus and their respective *S. zooepidemicus* orthologues *szo18890* and *szo18970*

Yellow bars represent orthology between the four inter-related orthologues of interest between *S. equi* and *S. zooepidemicus*. Red bars represent orthology between *S. equi* and *S. zooepidemicus*. The key *S. equi* paralogue is highlighted in bold.

3.3. PCR of *seq2180* and *seq2190*

Genomic DNA was extracted from cultures of *S. equi* 4047 and *S. zooepidemicus* H70 and used as template DNA for necessary amplifications (Figure 3.6). Due to an identical DNA 3'-terminal region between *seq2190* and its homologue *seq2180* it was necessary to perform a nested PCR with complementary reverse sites downstream from the 3'-terminal target to obtain the correctly amplified products *seq2190* and *seq2180*. Using primer pairs; Full2190_N_F,-Const_2190_R and Full_2180_N_F,-Const_2180_R two PCR fragments were amplified from the extracted genomic DNA of *S. equi* (See figure 3.7 and 3.8). Both primer sets amplified the correct nested construct yielding products of size consistent with the predicted length of 1900 bp and 1767 bp in length respectively. The products obtained were then further amplified to obtain only the mature target protein, removing the predicted N-terminal signal peptide sequence.

Using the primer pairs; Full2190_N_F,-Full2180/2190_E_R (a universal reverse primer for amplifications of both SEQ2190 and SEQ2180) and Full2180_N_F,-Full2180/2190_E_R two PCR fragments were amplified. The products yielded were of size consistent with the predicted lengths of 1884 bp and 1750 bp respectively (Figures 3.9 and 3.10). For investigation into the unique N-terminal sequence identified in SEQ2190, a further product was amplified consisting of only the variable *seq2190* region (truncated *seq2190*, encoding SEQ2190 N). This product was amplified from the nested *seq2190* amplimer using the primer set SEQFull2190_N_F,-Part_2190_E_R. The product obtained was shown to be of size consistent with its predicted length of 579 bp (Figure 3.11).

Notably this primer set also amplified additional larger PCR products; these were considered artefacts and not investigated further as no homologues of the target gene product were present in the *seq2190*-containing amplimer used as template.

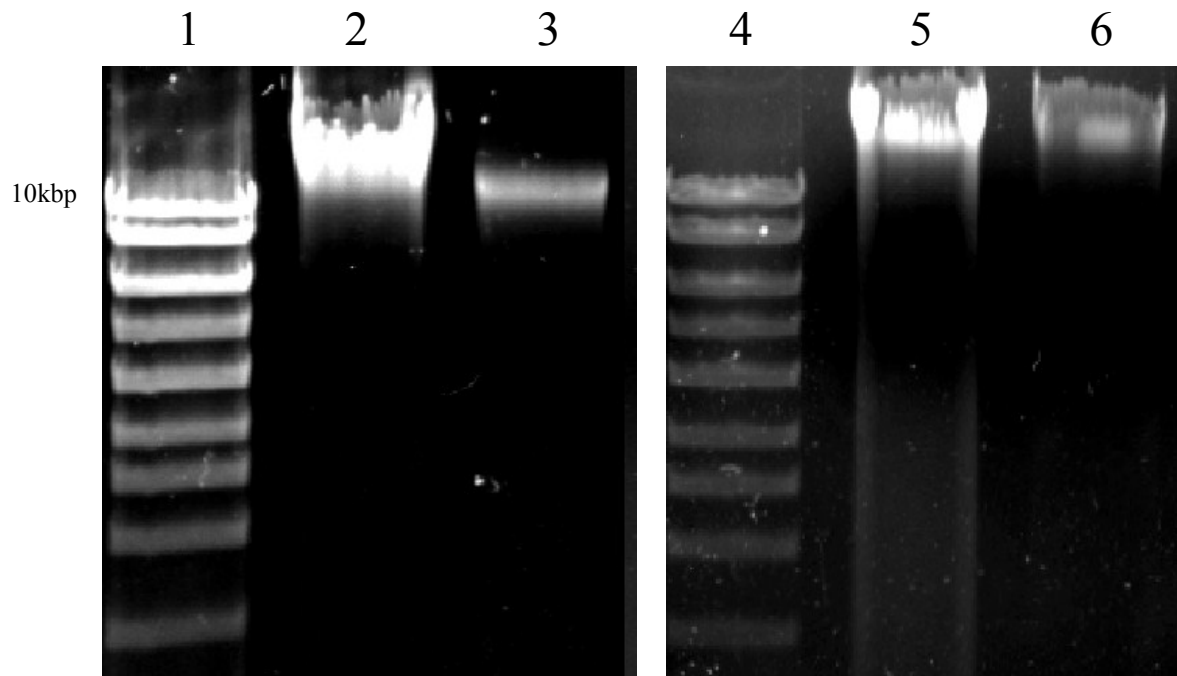


Figure 3.6: Genomic DNA extraction of *S. equi* 4047 and *S. zooepidemicus* H70.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 $\mu\text{g/ml}$). Lanes were as follows: lane 1, 10 kbp size standard; lane 2, extracted *S. equi* genomic DNA; lane 3, extracted *S. equi* genomic DNA (biological replicate); lane 4, 10 kbp size standard; lane 5, extracted *S. zooepidemicus* genomic DNA; lane 6, extracted *S. zooepidemicus* genomic DNA (biological replicate).

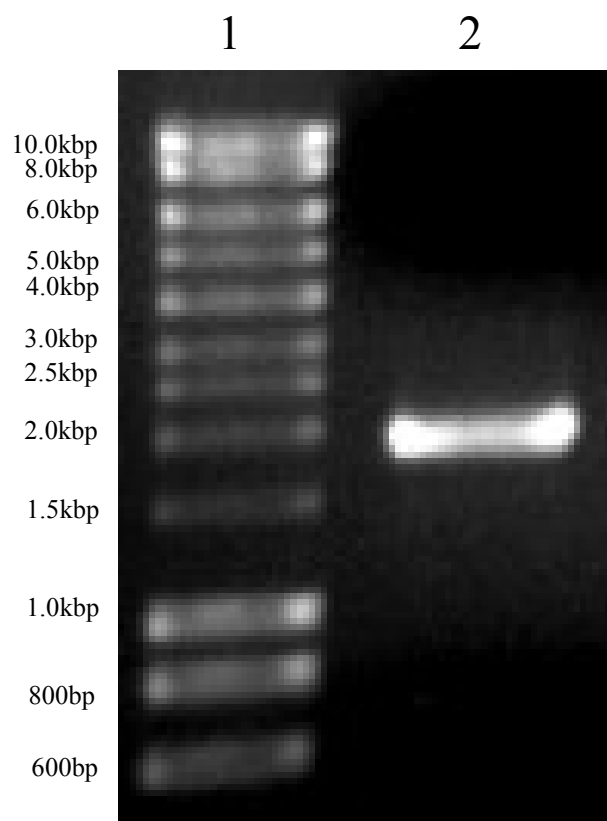


Figure 3.7: PCR amplification of *seq2190* and 3' extension.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were as follows: lane 1, 10 kbp size standard; lane 2, amplified product.

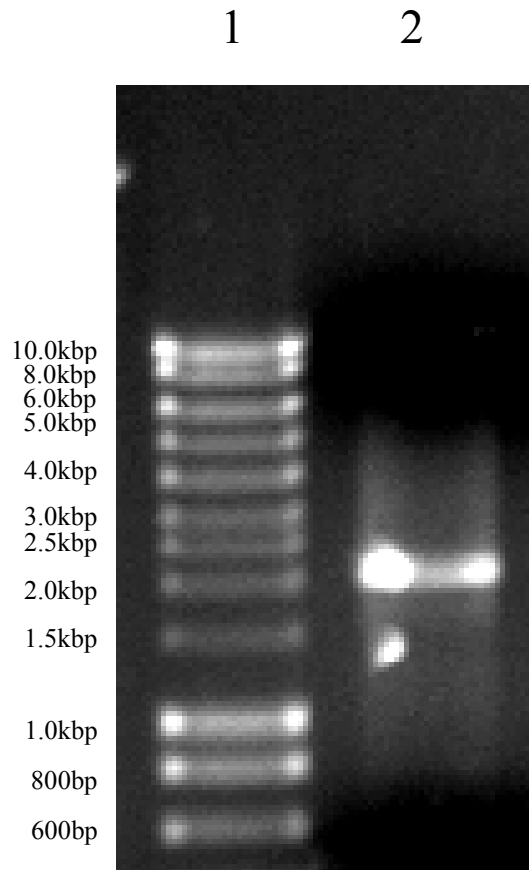


Figure 3.8: PCR amplification of *seq2180* and 3' extension.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standard; lane 2, amplified product.

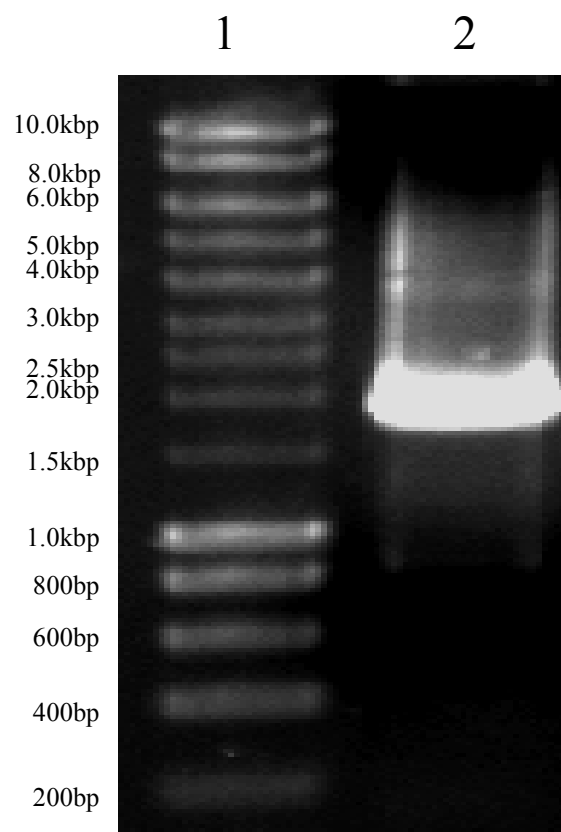


Figure 3.9: PCR amplification of *seq2190*.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standard; lane 2, amplified product.

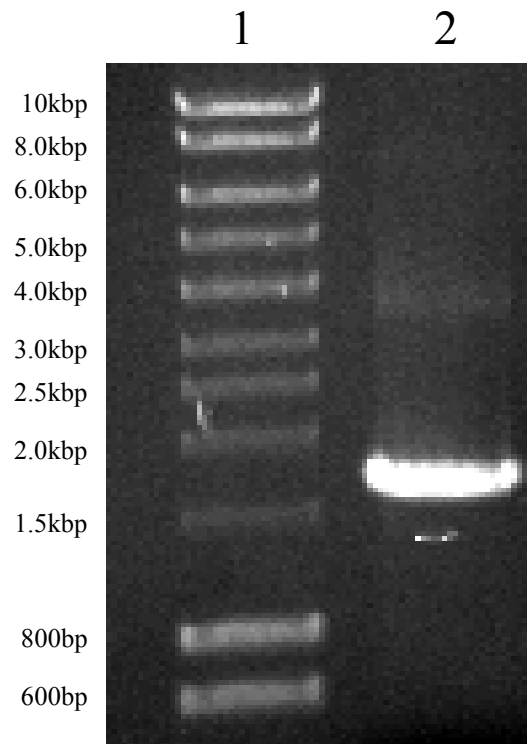


Figure 3.10: PCR amplification of *seq2180*.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standard; lane 2, amplified product.

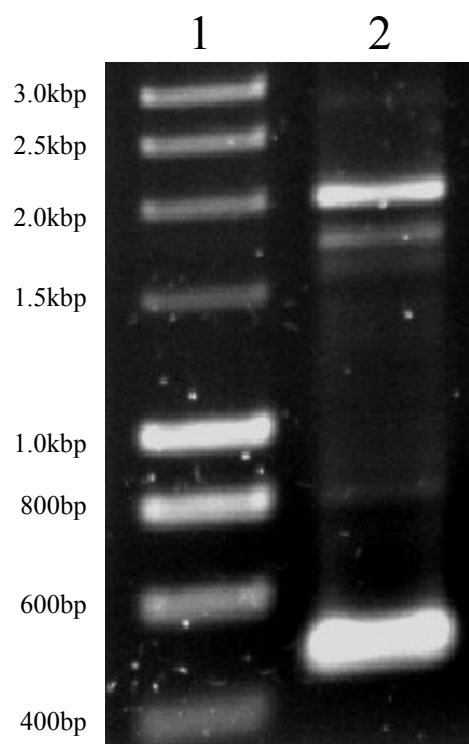


Figure 3.11: PCR amplification of truncated *seq2190* encoding SEQ2190 N.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standards; lane 2, amplified product.

3.4. Cloning of *seq2180* and *seq2190* (full length and truncated forms)

The amplified DNA products were initially cloned into pCR[®]-Blunt vector, yielding plasmids pb2190, pb2180 and pb2190 N (Table 2.5). These plasmids were subsequently digested with appropriate restriction endonucleases to verify product integration. Cloning vectors (target gene in pCR-Blunt) were then digested with appropriate endonucleases, purified and cloned into similarly restriction digested and purified pET 28a vector. This generated plasmids p2190, p2180 and p2190 N (Table 2.5) that encoded the proteins of interest with a His₆ tag at the N-terminus of the mature protein.

Confirmation of the correct plasmid construction was undertaken by screening of plasmids extracted from candidate positive clones with appropriate restriction endonucleases. Following digestion, a successful clone was indicated by the visualisation of a product of the expected target gene size and the remaining restriction endonuclease cut pET 28a vector, 5.4 kbp in size. The expected sizes for the products of the inserts in pb2180/p2180, pb2190/p2190 and pb2190/p2190 N are 1767 bp, 1884 bp and 579 bp as can be clearly seen in figures 3.12, 3.13 and 3.14. Each putative correct construct was sequenced for further confirmation.

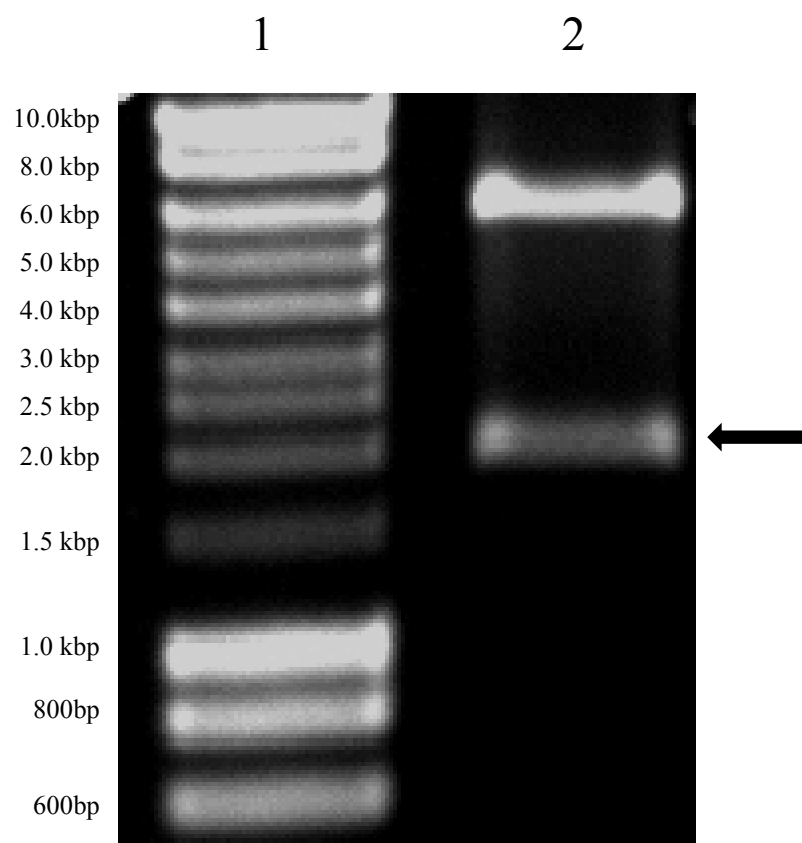


Figure 3.12: Restriction endonuclease digestion of p2190 plasmid.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 $\mu\text{g/ml}$). Lanes were loaded as follows: Lane 1, 10 kbp size standard; lane 2, p2190 post treatment with *NdeI* and *XhoI* restriction endonucleases. Released *seq2190* product (expected size 1884 bp) is indicated with an arrow. The upper band is the pET28a vector.

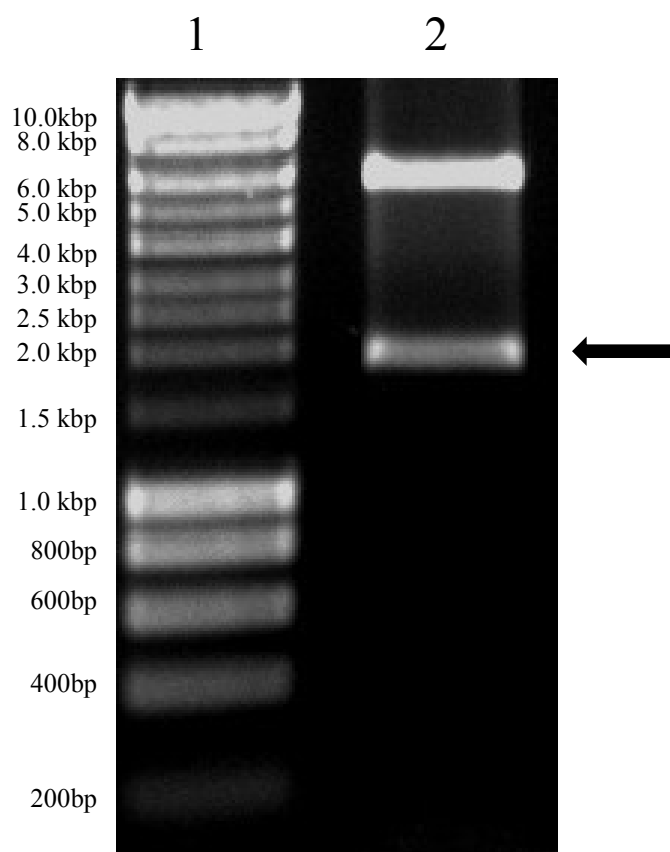


Figure 3.13: Restriction endonuclease digestion of p2180 plasmid.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 $\mu\text{g/ml}$). Lanes were loaded as follows: Lane 1, 10kbp size standard; lane 2, p2180 post treatment with *NdeI* and *EcoRI* restriction endonucleases. Released *seq2180* product (expected size 1767 bp) is indicated with an arrow. The upper band is the pET28a vector.

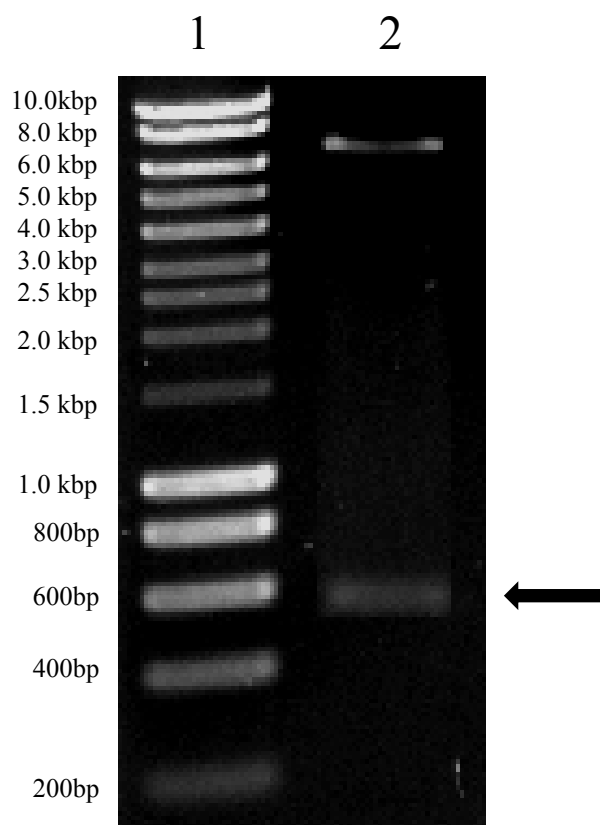


Figure 3.14: Restriction endonuclease digestion of p2190 N plasmid.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standard; lane 2, p2190N post treatment with *NdeI* and *EcoRI* restriction endonucleases. Released truncated *seq2190* product (expected size 579 bp) is indicated with an arrow. The upper band is the pET 28a vector.

3.5. Protein expression

Plasmids p2190 p2180 and p2190 N, encoding SEQ2190, SEQ2180 and the SEQ2190 N respectively, were chemically transformed into a range of *E. coli* expression strains (Table 2.1) and grown at 37°C to mid-exponential phase. Initially, the expression of the recombinant proteins from these strains was induced by the addition of 1 mM IPTG (final concentration) and the cultures were incubated overnight. Whole cell protein extractions were taken both at 3 hours and 16 hours post induction. A temperature of 20°C was initially chosen as reduced temperature is a parameter known to limit the *in vivo* aggregation of recombinant proteins (Schein 1989). The whole cell protein extracts obtained were compared to a whole cell extract of *E. coli* BL21 (DE3) with no transformed expression plasmid present

Whole cell protein extracts visualised by SDS-PAGE showed expression of the target SEQ2190 protein from plasmid p2190 best in two specific strains of *E. coli*, BL21 and B834 (DE3) (Figure 3.15). However further analysis of expression of SEQ2190 in *E. coli* Tuner™ indicated a small visual increase in expressed protein and less variability regarding whole cell harvesting between 3 and 16 hours (Figure 3.16). The expressed protein visualised was consistent in size to the predicted protein size of His₆-Tagged mature SEQ2190 (68.7 kDa).

Furthermore, plasmids p2180 and p2190 were also found to have favourable levels of cloned protein expression in *E. coli* Tuner™, expressing proteins consistent in size to the predicted size of each His₆- Tagged mature protein, 64.2 kDa and 24.4 kDa respectively (Figures 3.17 and 3.18). Expression of proteins were found to be favourable with 1.0 mM IPTG (final concentration) in cultures agitated at 100rpm at 37°C overnight (16 hours).

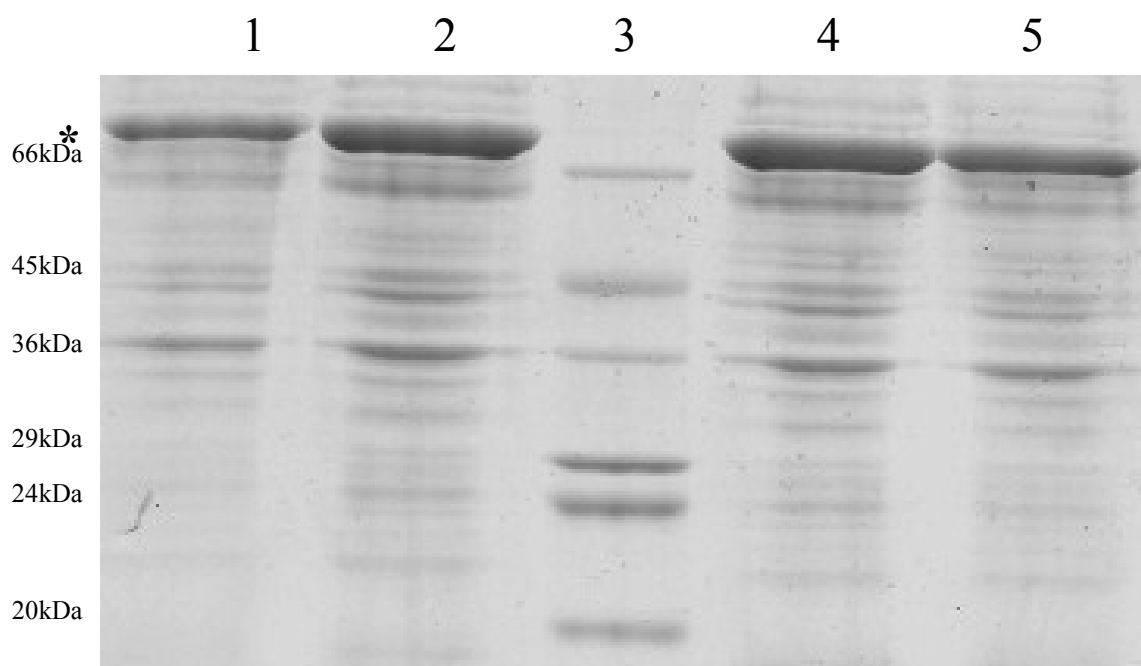


Figure 3.15: 15% SDS PAGE gel demonstrating SEQ2190 expression trials.

SEQ2190 expression was induced by 1 mM IPTG and growth at 20°C. The expected position of the SEQ2190 protein is marked with an asterisk. Lanes were loaded as follows: Lane 1, *E. coli* BL21 (DE3) p2190, 3 hour induction; lane 2, *E. coli* BL21 (DE3) p2190, 16 hour induction; lane 3, molecular weight markers (66, 45, 36, 29, 24, 20 kDa); lane 4, *E. coli* B834 (DE3) p2190, 3 hour induction; lane 5, *E. coli* B834 (DE3), 16 hour induction.

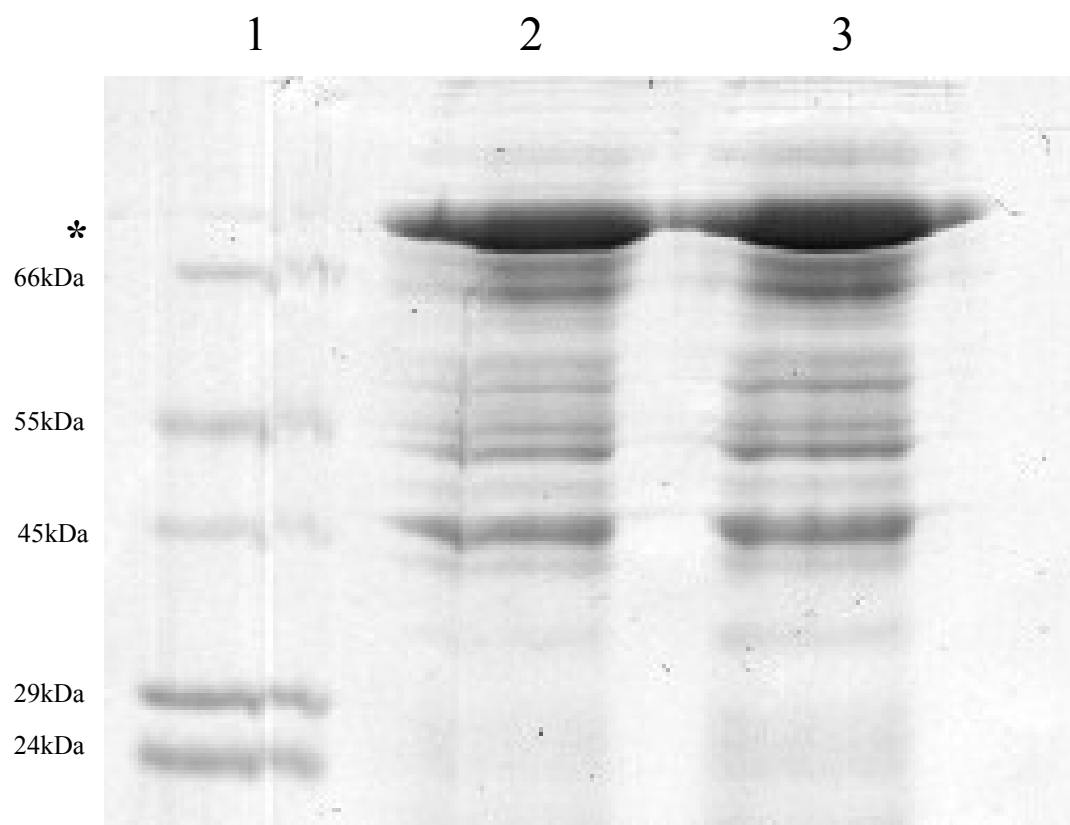


Figure 3.16: 15% SDS PAGE gel demonstrating SEQ2190 expression

SEQ2190 expression was induced by 1 mM IPTG and growth at 20°C. The expected position of the SEQ2190 protein is marked with an asterisk. Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24 kDa); lane 2, *E. coli* Tuner® (DE3) p2190, 3 hour induction; lane 3, *E. coli* Tuner® (DE3) p2190, 16 hour induction.

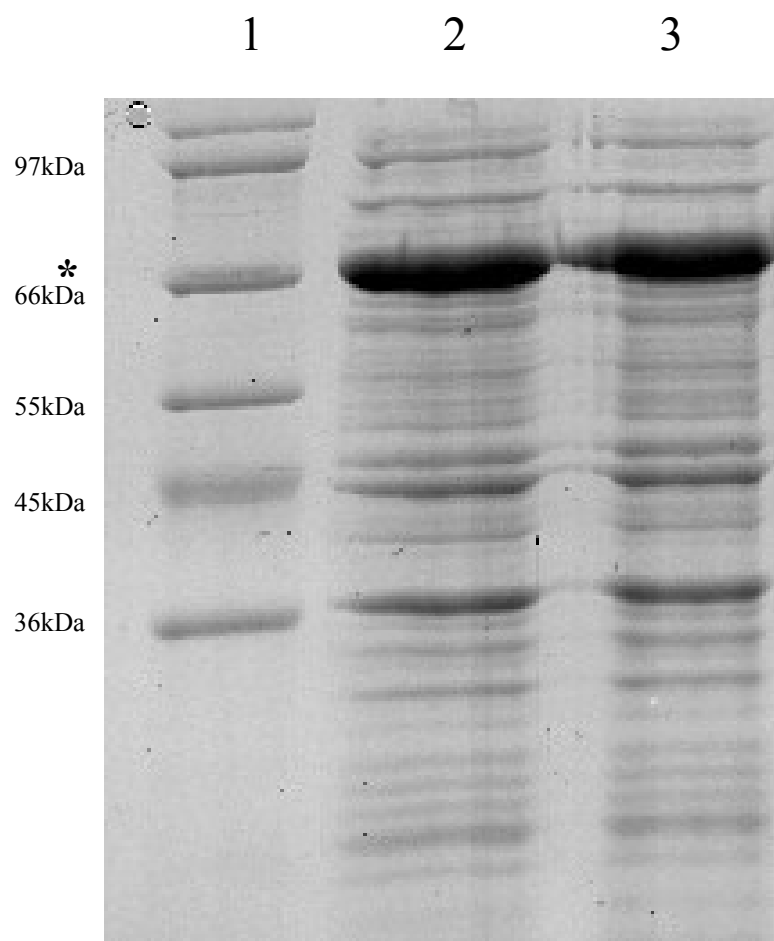


Figure 3.17: 15% SDS PAGE gel demonstrating SEQ2180 expression.

SEQ2180 expression was induced by 1 mM IPTG and growth at 20°C. The expected position of the SEQ2180 protein is marked with an asterisk. Lanes were loaded as follows: Lane 1, molecular weight markers (97, 66, 55, 36, kDa); lane 2,- *E. coli* Tuner® (DE3) p2180, 3 hour induction; lane 3,- *E. coli* Tuner® (DE3) p2180, 16 hour induction.

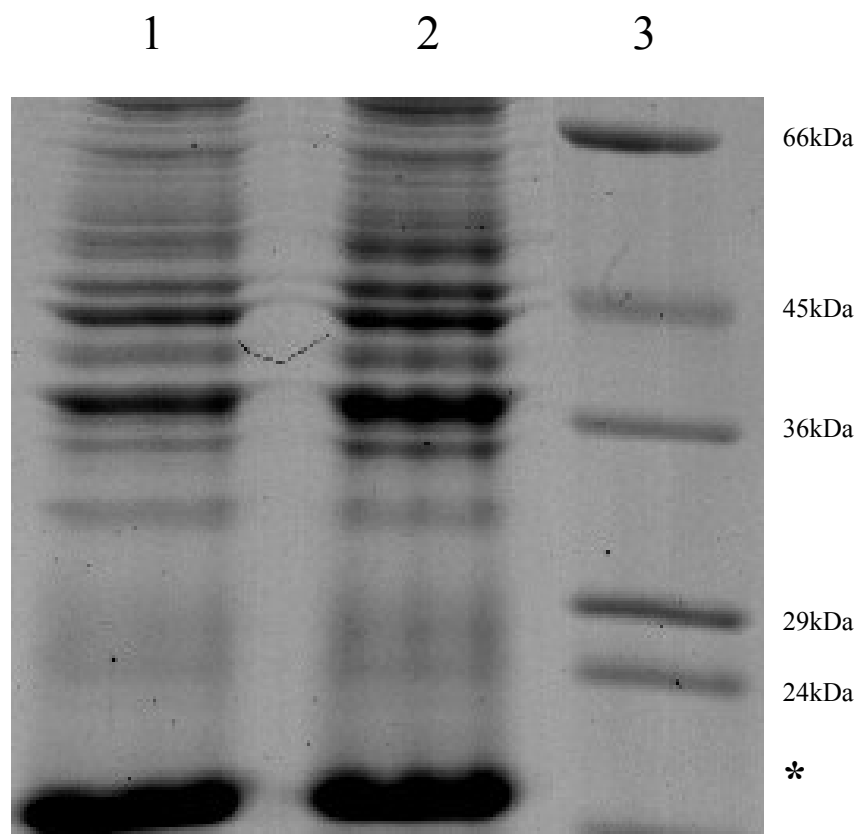


Figure 3.18: 15% SDS PAGE gel demonstrating SEQ2190 N expression.

SEQ2190N expression was induced by 1 mM IPTG and growth at 20°C. The expected position of the SEQ2190 N truncated protein is marked with an asterisk. . Lanes were loaded as follows: Lane 1, *E. coli* Tuner® (DE3) p2190N, 3 hour induction; lane 2, - *E. coli* Tuner® (DE3) p2190N, 16 hour induction; lane 3, molecular weight markers (66, 45, 36, 29, 24 kDa).

3.6. Protein purification

All His₆ –tagged proteins were expressed from cultures in 4 l of LB media and purified from cell free extracts by IMAC using a chelating Sepharose column and an automated gradient elution technique (section 2.13.4). Identification of proteins present in different fractions was indicated by U.V absorption. Fractions containing expressed SEQ2190, SEQ2180 and SEQ2190 N were checked for purity by SDS-PAGE (Figures 3.19, 3.20 and 3.21 respectively) and concentrated. Proteins SEQ2180, SEQ2190 and SEQ2190 N were then further purified by FPLC, monitored by U.V absorption and checked for purity by SDS-PAGE, concentrated and quantified (Figures 3.22, 3.23 and 3.24 respectively). The quantities of purified SEQ2180, SEQ2190 and SEQ2190 N obtained were varied depending on the protein i.e. ~1.25 mg/L, ~1.75 mg/L and ~2.5 mg/L respectively.

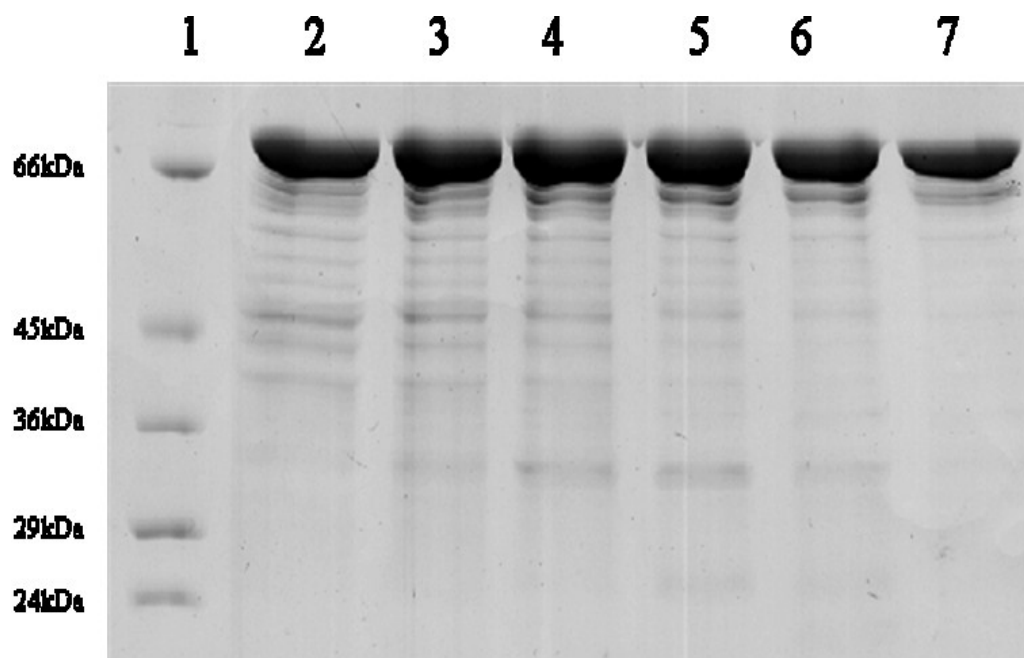


Figure 3.19: 15% SDS PAGE of IMAC purified fractions of SEQ2190.

Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 20 kDa); lanes 2-7, individual IMAC fractions of purified SEQ2190 (fractions 48, 46, 44, 42, 40, 38).

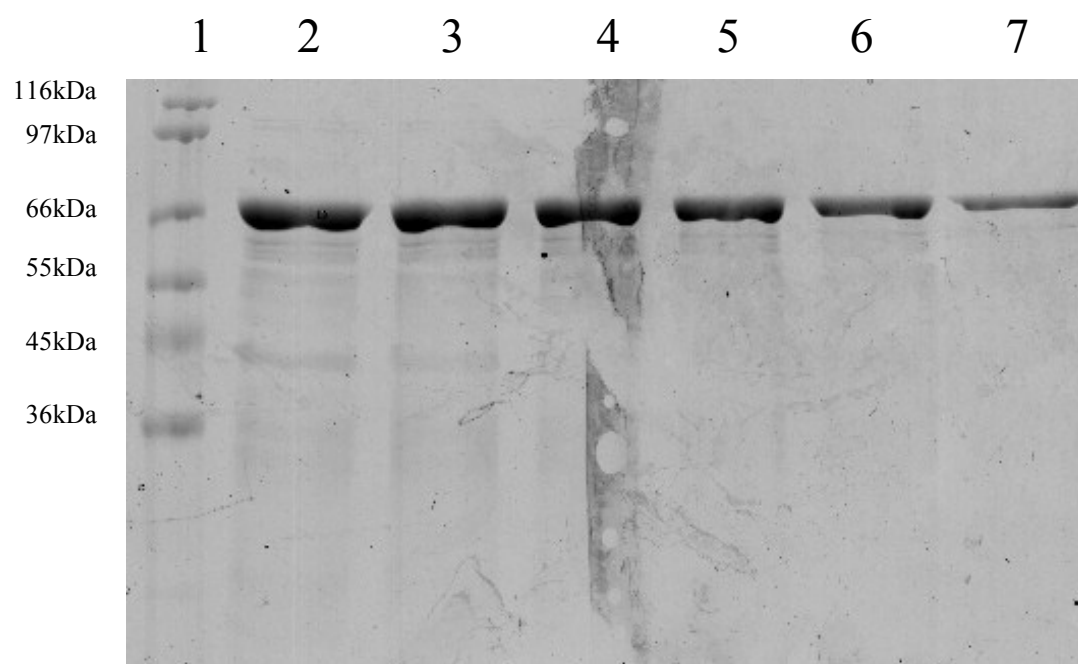


Figure 3.20: 15% SDS PAGE gel of IMAC purified fractions of SEQ2180.

Lanes were loaded as follows: Lane 1, molecular weight markers (116, 97, 66, 55, 45, 36 kDa); lanes 2-7, individual IMAC fractions of purified SEQ2180 (fractions 47, 49, 51, 53, 55, 57,)

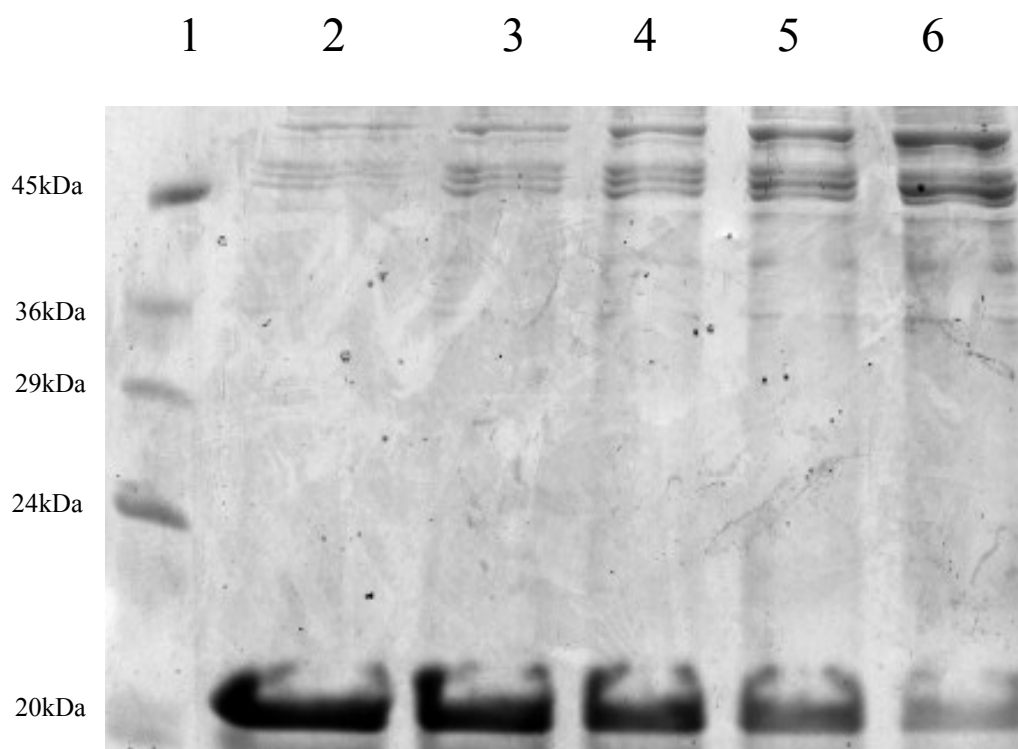


Figure 3.21: 15% SDS PAGE of IMAC purified fractions of SEQ2190 N.

Lanes were loaded as follows: Lane 1, molecular weight markers (45, 36, 29, 24, 20 kDa); lanes 2-6, individual fractions of purified SEQ2190_N (fractions 45, 43, 41, 39, 37).

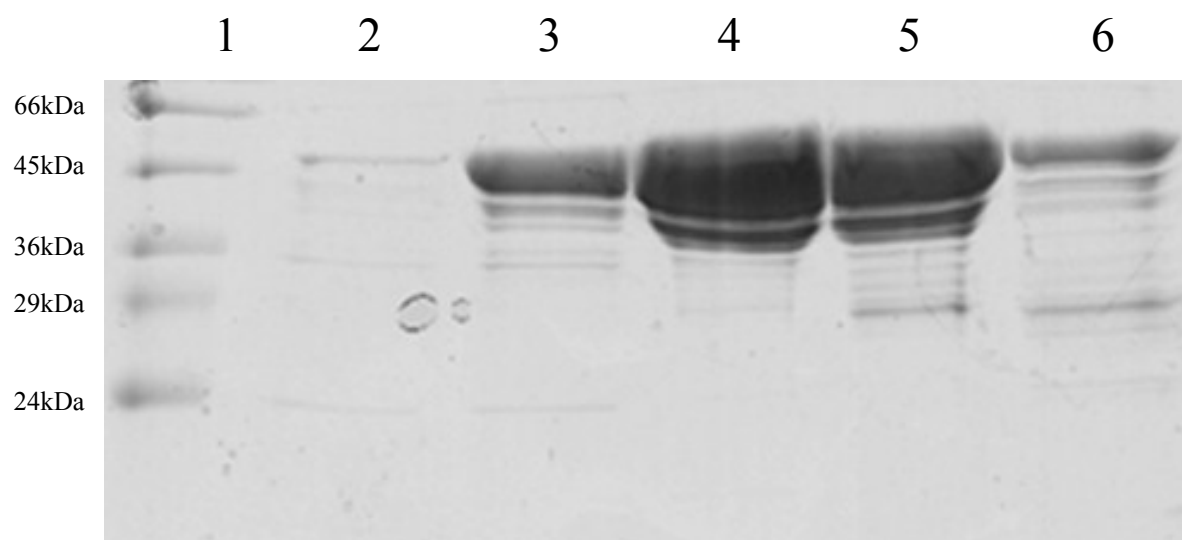


Figure 3.22: 15% SDS PAGE gel showing FPLC purified fractions of SEQ2190.

Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24 kDa); lanes 2-6, individual fractions of purified SEQ2190_N (Fractions 15, 14, 13, 12, 11).

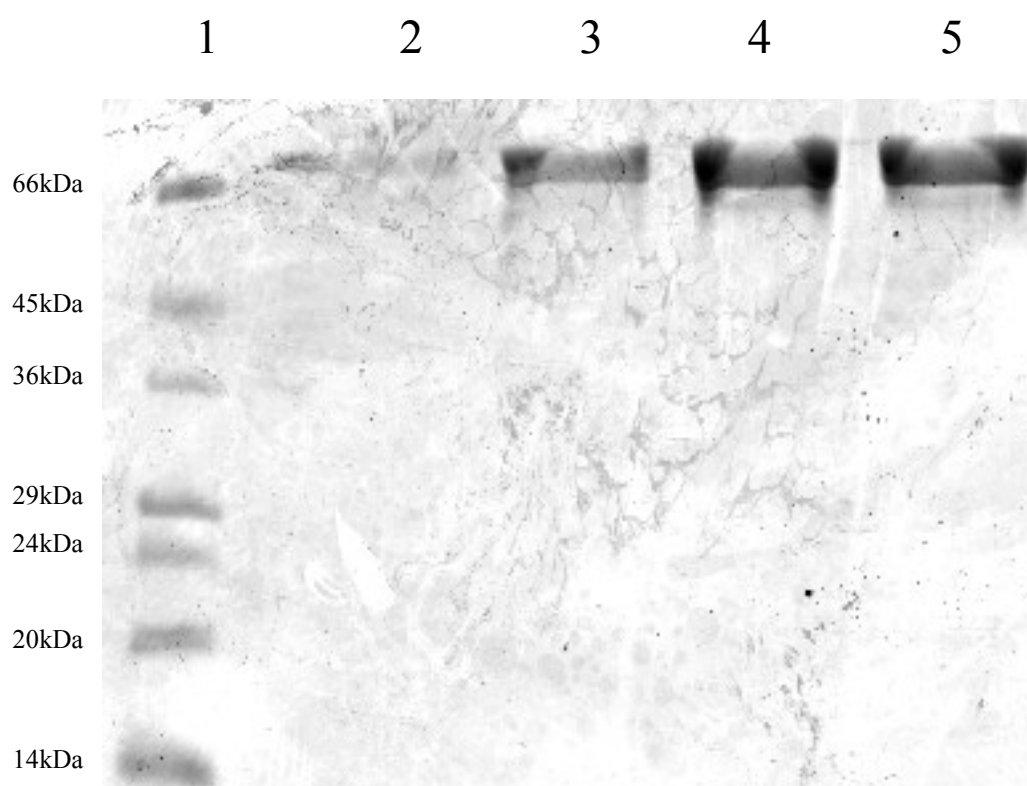


Figure 3.23: 15% SDS PAGE gel showing FPLC purified fractions of SEQ2180.

Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 20, 14 kDa); lanes 2-5, individual fractions of purified SEQ2180 (Fractions 12, 13, 14, 15).

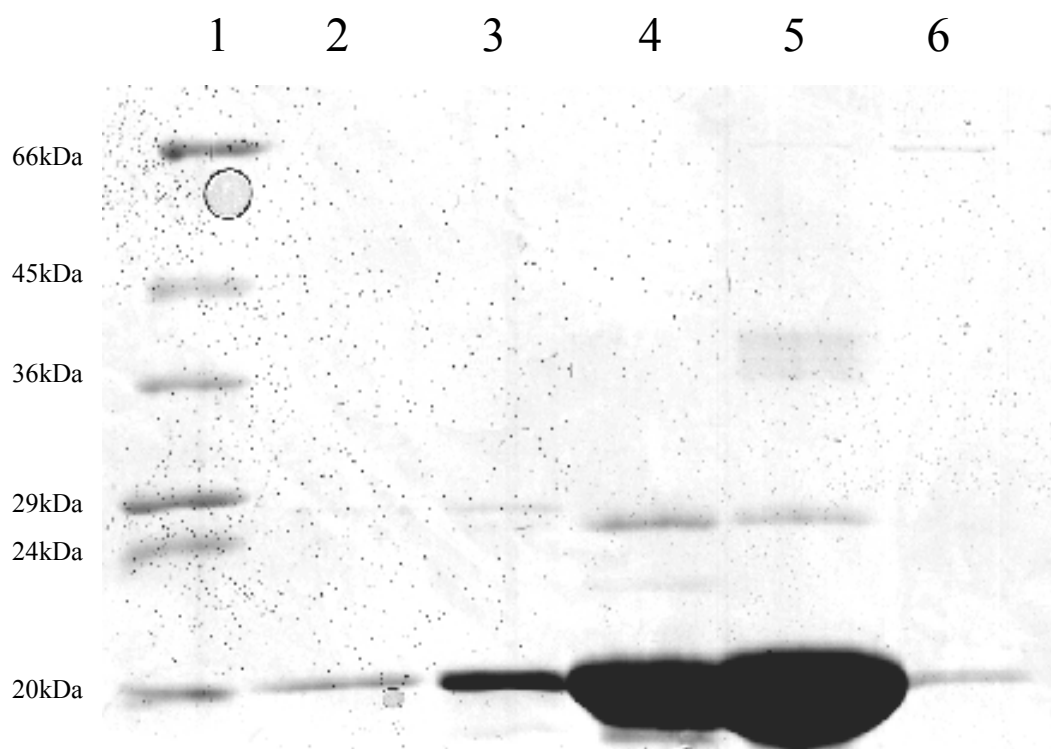


Figure 3.24: 15% SDS PAGE gel of FPLC purified fractions of SEQ2190 N

Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 20 kDa); lanes 2-6, individual fractions of purified SEQ2190 N (Fractions 25, 23, 21, 19, 17, 15).

3.7. Preperation of crystals for X-ray crystallography

A number of different crystallisation screens were used to attempt to crystallise SEQ2180, SEQ2190 and the truncated SEQ2190 N proteins. These included PEG/Ion, Hampton Screen (HS) I and II, PEG/Anion, PEG/Cation, Newcastle and Index screens (See appendix III for formulations).

Crystals of the SEQ2190 proteins were recovered from the PEG/Ion screens 35 and 3 (Figures 3.25 and 3.26) but unfortunately neither of these crystals was suitable for diffraction. Unfortunately no other crystals were observed for any other proteins despite extensive screening with the screen library available.



Figure 3.25: Crystal of SEQ2190

Ratio of protein and buffer added was 1:1. The condition tested that yielded the crystal was 0.2M ammonium sulphate, 20% w/v polyethylene glycol 3350, pH 6.0.

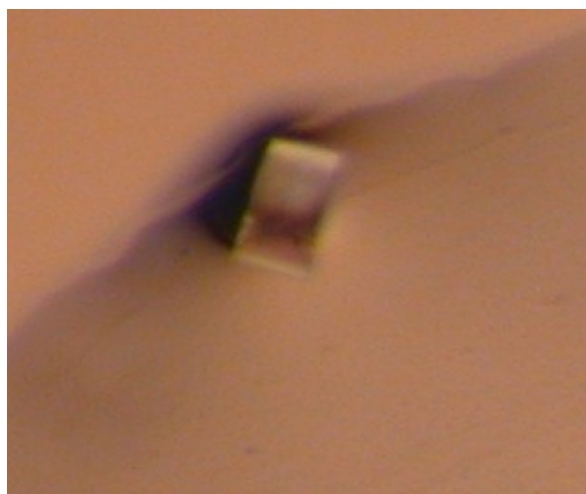


Figure 3.26: Crystal of SEQ2190.

Ratio of protein and buffer added was 1:1. The condition tested that yielded the crystal was 0.2M ammonium fluoride, 20% w/v polyethylene glycol 3,350, pH 6.2.

3.8. Discussion and future work

The data provided in this study has demonstrated the production of three useful plasmid constructs. Furthermore it has been demonstrated that useful quantities of purified SEQ2180, SEQ2190 and the truncated SEQ2190 N protein could be produced using these constructs for further research.

Bioinformatic analysis has indicated that SEQ2190 represents a protein of three distinct domains with a significant N-terminal domain, unique to *S. equi* 4047 of particular note. This protein is accompanied by a structurally similar but N-terminal sequence variable paralogue encoded by *seq2180*. Moreover, an N-terminally deficient pseudogene homologue of *seq2190* is also present in *S. zooepidemicus* (Figure 3.3). It can be speculated that an ancestral strain of *S. equi* and *S. zooepidemicus* may have possessed only one of these paralogues, which then duplicated prior to the sub-species divergence event. Furthermore it can be speculated that an earlier strain of *S. zooepidemicus* may have retained a functional copy of SEQ2190 which has since become lost. The previous work undertaken by collaborators at the Animal Health Trust indicated that SEQ2190 was conserved in all tested strains of *S. equi* and the loss of the N-terminal portion of the *S. zooepidemicus* homologue was consistent amongst strains.

Whilst it was indicated that the *S. zooepidemicus seq2190* pseudogene homologue was also conserved it remains to be determined what level of variability, if any, is present upstream from the point of mutation that has isolated the coding region for the C-terminal section of the protein as pseudogene. Although also attempted in this study, the lack of loci synergism amongst paralogous *S. zooepidemicus* strains in this region has meant that typical

chain-termination methods of sequencing are not possible. It is however expected that a high-throughput genome sequencing methods and the publication of further whole genome sequences of *S. zooepidemicus* strains will provide further insight into the *seq2190* pseudogene, the function of which was presumably lost as a result of a truncation event.

It was also the aim of this study to deduce a 3-dimensional structure of SEQ2190 and SEQ2180 following successful diffraction of respective protein crystals. Although work was undertaken to achieve this and some crystals of SEQ2190 produced it is regrettable that in the time frame of this study that this aim was not completed. However a proposed structure can be predicted following the computational analysis undertaken in this study (Figure 3.27). This proposed structure is primarily a helical stalk structure and a β -strand portion representing of the previously discussed unique N-terminal region. The extensively alpha helical structure is likely to be necessary to project the N-terminal domain up to the cell surface and through the hyaluronic acid capsule. It is this β -strand region that is exclusive to the mature SEQ2190 proteins and hence the potentially significant putative host ligand interactive portion of each protein that may contribute to pathogenesis.

One such benefit from the availability of 3-Dimensional structures of SEQ2180 and SEQ2190 would be clearer insight into potential ligands or catalytic activities. As demonstrated, other than in *S. zooepidemicus*, SEQ2190 has only orthologues of low significance in *S. pyogenes* and these are also highly helical proteins projecting from the cell wall (Timoney, et al. 1997). Indeed the homologous regions are only located within the common helical regions. It was expected that the solving of a 3-dimensional structure would have provided the tools to enable more stringent structure database searches for potential ligand candidates for SEQ2190 and SEQ2180. It should also be noted that the degree of ligand specificity has not been established for either SEQ2190 or SEQ2180 nor has the

potential of possible co-receptor ligand binding. However the absence of a functional partner to SZO18890 in *S. zooepidemicus* perhaps argues against this possibility.

To provide insight into SEQ2190 as a potential virulence factor, it was also the initial aim of this study to produce $\Delta seq2190$, $\Delta seq2180$ and $\Delta seq2190$ *N* strains of *S. equi* based on methodology described by Hamilton and colleagues (Hamilton, et al. 2006). However at the time of writing this has not yet been successful. It was expected that the production of these mutant strains would have enabled further study in a mouse model of strangles (Chanter, et al. 1995) and investigation of changes (if any) in the surface epithelial morphology of air interface organ cultures (Jackson, et al. 1996) exposed to these mutant *S. equi* strains.

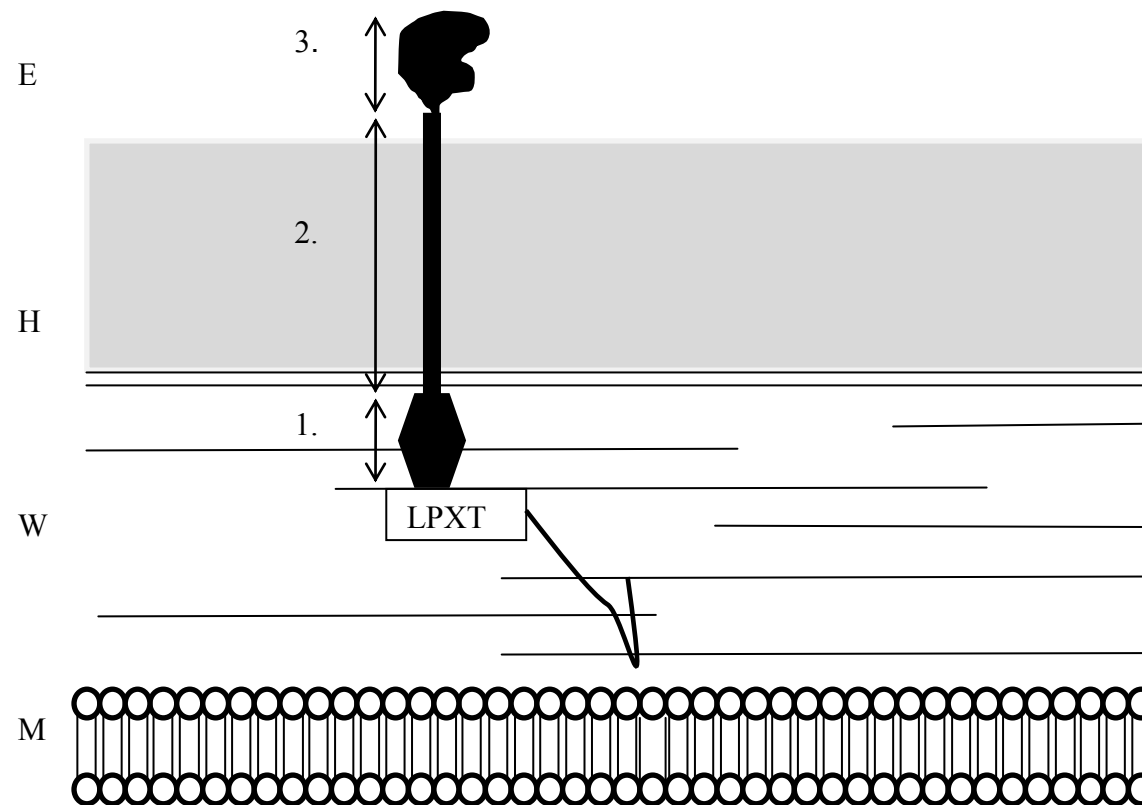


Figure 3.27: Proposed predicted secondary structure model of the mature SEQ2190 protein with three distinct structural domains.

(1) C-terminal region of predominantly protein coil. (2) Core region of long extended α -helices (3) N-terminal region of β -strands, protein coil and α -helices; W, cell wall; H, Hylauroic acid capsule; E, Exposed bacterial surface

4. Investigation into the secreted Phospholipase A₂ enzyme, SlaA of *S. equi*.

4.1. Introduction

As described in section 1.8, PL represents an interesting class of multifunctional and diverse enzymes with particular attention given to sPLA₂ enzymes. With regards to the streptococci, the ability of sPLA₂ to result in the cleavage products, lysophospholipid and free fatty acid (crucially AA) represents an interesting prospect as a potential virulence factor. The work of Sitkiewicz et al (2006) in particular has already provided clear evidence for potency of sPLA₂ as a virulence factor with demonstrated effects of the *S. pyogenes* sPLA₂, SlaA (here in referred to as SpSlaA) on eukaryotic cells.

Since *S. pyogenes* shares many virulence factors with *S. equi* and *S. zooepidemicus* reasonable speculation can be given that the presence of SlaA in *S. equi* would possess similar potency as a virulence factor.

The presence of SlaA however is not conserved amongst *S. equi* and *S. zooepidemicus* strains as shown by the work of Holden *et al* (2009) indicating that the presence of SlaA in some *S. equi* strains may have been acquired via mobile genetic elements similarly to its presence in the pathogenically more severe *S. pyogenes* M3 strains. SlaA has been detected in 26 of 26 tested strains of *S. equi* (crucially including the M-protein plus strain, CF32) but only detected in 44 of 140 tested strains of *S. zooepidemicus*, indicating the acquisition/retention of *slaA* in *S. equi* has a far greater biological significance than its acquisition/retention in *S. zooepidemicus*. The absence of SlaA in a large proportion of *S.*

zooepidemicus strains also implies a potential molecular difference that may relate to the clear pathogenic contrast shown between *S. equi* and *S. zooepidemicus* (see section 1.2).

Isolates from the early 20th century (*S. equi* strain ATCC 9527) have been shown to encode SlaA (Ikebe et al., 2002), predating the isolation of SpSlaA expressing *S. pyogenes* strains (Beres, et al. 2002) also indicates that it is possible that SpSlaA was acquired in *S. pyogenes* via phage-mediated horizontal transfer from a SlaA-containing strain of *S. equi* (Holden, et al. 2009; Ikebe, et al. 2002).

Currently it is not known if SlaA expressed by *S. equi* has the potential to trigger an inflammatory cascade via the cleaving of phospholipids, releasing AA. The degree of enzymatic activity expressed by SlaA and its paralogue SlaB in comparison with other similar sPLA₂ activities also remains undetermined.

According to current sPLA₂ nomenclature (Burke and Dennis 2009) , SlaA and SlaB may represent a new family of sPLA₂ enzymes with the molecular mass and the number of disulphide bonds present not consistent amongst any of the currently proposed groups.

Aims of the present study

The aims of the present study were:

- a) to clone and express SlaA identified in *S. equi* 4047 and SlaB identified in *S. equi* 4047 and *S. zooepidemicus* H70 in *E. coli*.
- b) to obtain purified forms of SlaA and SlaB.
- c) to assess the enzymatic properties of SlaA and SlaB and provide a comparison to previously characterised similar sPLA₂ enzymes.
- d) to investigate the ability of SlaA and SlaB to release AA from phospholipids.

4.2. Bioinformatic identification of SlaA and an SlaA homologue in *S. equi*.

The amino acid sequence of the putative sPLA₂, SlaA (SEQ0849) was obtained from UniProt in a FASTA format. The SEQ0849 sequence was analysed for identification of a signal peptide cleavage site by using the SignalP Hidden Markov Models and Neural Networks outputs, with a cleavage site predicted between position 27 and 28, with a probability of 1.0. A canonical lipoprotein ‘lipobox’ (LAAC13) is noted to be present in the signal peptide, but due to the atypical shortness of the preceding H- region, SlaA is not predicted to be a putative lipoprotein. A BLAST search of the translated *S. equi* genome database for homologues to SlaA revealed an amino acid sequence with 80% identity (130/191 amino acids), annotated as SlaB (SEQ2155). The amino acid sequence of SEQ2155 also differed from SEQ0849 in that the signal peptide cleavage site was instead predicted between position 26 and 27. Notably, the predicted non-functional lipobox in the SlaA signal peptide is not conserved in SlaB.

A BLAST search was undertaken for both SEQ0849 and SEQ2155 of the *S. equi* sub sp. *zooepidemicus* H70 genome project (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_zooepidemicus). Notably, no orthologue to SlaA was found in *S. zooepidemicus* H70. However, an identifiable orthologue of SEQ2155 was identified in *S. zooepidemicus* H70 showing 90.0% identity (172/190 amino acids, SZO18670), (Figure 4.1).

Database searches of homologues to the amino acid sequence of the mature SlaA (SEQ0849) protein using BLAST revealed a large number of proteins with significant similarities, and some proteins with low similarities but of biological significance (Table 4.1). Of particular interest, is the SpSpaA protein investigated by Nagiec *et al.* (2004) and identified as a phage encoded protein in the genomes of *S. pyogenes* strains MGAS315

(Beres et al 2002) and N1H1 (Ikebe et al. 2002). This orthologue was shown to have 99% identity (162/164 amino acids) to *S. equi* SlaA (Figure 4.2).

Database searches of protein families also revealed that SlaA and SlaB were likely to be representatives of the sPLA₂ family of phospholipase enzymes. BLAST analyses noted a significant match to the Pfam family, Phospholip_A2_1 (PF00068), Phospholipase A2 (Figure 4.3) and also noted less significant matches to the Pfam family PLA2G12 (PF06951, Group XII secretory phospholipase A2 precursor) and Phospholip_A2_2 (PF05826, Phospholipase A2). As indicated in Figure 4.1, the putative Ca²⁺ loop consensus sequence region and active site consensus region was identified based on putative regions defined in SpSlaA (Nagiec, et al. 2004).

Table 4.1: Representative orthologues of SEQ0849.

<u>Accession</u>	<u>Annotation</u>	<u>Organism</u>	<u>Identities</u>	<u>Positives</u>	<u>Gaps</u>	<u>E number</u>
gb AY050245.2	Putative methionine sulfoxide reductase gene	Streptococcus pyogenes strains NIH1 and MGAS315	162/164 (99%)	164/164 (100%)	0/164 (0%)	1.00E-92
gb GQ923934.1	Phospholipase A2 (slaA gene)	Streptococcus dysgalactiae subsp. equisimilis strain B107	162/164 (99%)	164/164 (100%)	0/164 (0%)	3.00E-92
gb CP001709.1	Plasmid pAPRE01	Anaerococcus prevotii DSM 20548	72/159 (45%)	101/159 (64%)	5/159 (3%)	5.00E-31
gb CP002618.1	Hypotential protein	Lactobacillus casei BD-II	38/100 (38%)	52/100 (52%)	14/100 (14%)	3.00E-07
gb CP001001.1	Conserved hypothetical protein	Methylobacterium radiotolerans JCM 2831	25/50 (50%)	30/50 (60%)	5/50 (10%)	7.00E-05


```

SlaA      MKKVINTILLAACFLFFLGNFTTNVLAEGINDKIENGTESDISFQNGELLKNYLILEGER 60
SpSlaA    MKKVINTILLAACFLFFLGNFTTNVLAEGINDKMENGTESDISFQNGELLKNYLILEGER 60
          *****:*****

SlaA      VYFDYDRATQDKVSDDVLEMGMLEVAISKDYSEKTFTPDKYFKASWPIHGNYCGPGHNGN 120
SpSlaA    VYFDYDRATQDKVSDDVLEMGMLEVAISKDYSEKTFTPDKYFKASWPVHGNYCGPGHNGN 120
          *****:*****

SlaA      NFTLPVVDVLDQGCQNHDSCYKWGAGIGANCECNRQLVNYIKVYRRWMPANVLGVADAIR 180
SpSlaA    NFTLPVVDVLDQGCQNHDSCYKWGAGIGANCECNRQLVNYIKVYRRWMPANVLGVADAIR 180
          *****

SlaA      VYFETVGSIGC 191
SpSlaA    VYFETVGSIGC 191
          *****

```

Figure 4.2: Pairwise alignment of SlaA and the SpSlaA protein encoded by temperate phage of *S. pyogenes* strains.

Alignment was performed using ClustalW2. * indicates identical amino acids, : indicates similar amino acids. The two variant amino acid positions are underlined.

#HMM	EYAEYGCYCGAGGSGTP.....KDELDRCKVHDECYEEAEK..KGCKPKLKTYSY
#MATCH	++ +G yCg g +G+ +D lD+ C+ hd+Cy+ +c+ + + +y
#PP	566789*****8433334799*****965443456665555544
#SEQ	SWPIHGNYCGPGHNGNftlpvVDVLDQGCQNHDSCYKWGAGigANCECNRQLVNY

Figure 4.3 Identification of Pfam family.

The conserved PLA₂ domain was identified between position 78 and 133 of the mature SlaA (SEQ0849) protein with an Expect value of 0.00021, indicating that the likelihood of a correct match is high. #HMM refers to the consensus of the hidden markov models (HMM), #MATCH refers to the match between the Pfam family and the amino acid sequence of SEQ0849, #PP refers to the posterior probability and #SEQ refers to the matched portion of SEQ0849.

4.3. PCR of *seq0849* and *seq2155*

Using the designed primers pairs- SEQ0849f-SEQ0849r and SEQ2155f-SEQ2155r (Table 2.4), two PCR fragments were amplified from the extracted genomic DNA of *S. equi* (Figure 3.4). Both primer sets amplified only the mature protein and removed the predicted N-terminal signal peptide sequence and the reverse primer of each primer set contained its complementary base sequence downstream from the target gene. Primer set SEQ0849f and SEQ0849r yielded a product of size consistent with its predicted length of 717bp (Figure 4.4) and primer set SEQ2155f and SEQ2155r yielded a product of size consistent with its predicted length of 779bp in length (Figure 4.5). Notably these primers also amplified additional larger PCR products which were not investigated further as BLAST searches strongly suggest that SEQ0849 and SEQ2155 are the only putative sPLA₂ paralogues encoded in the *S. equi* 4047 genome (Figure 4.5).

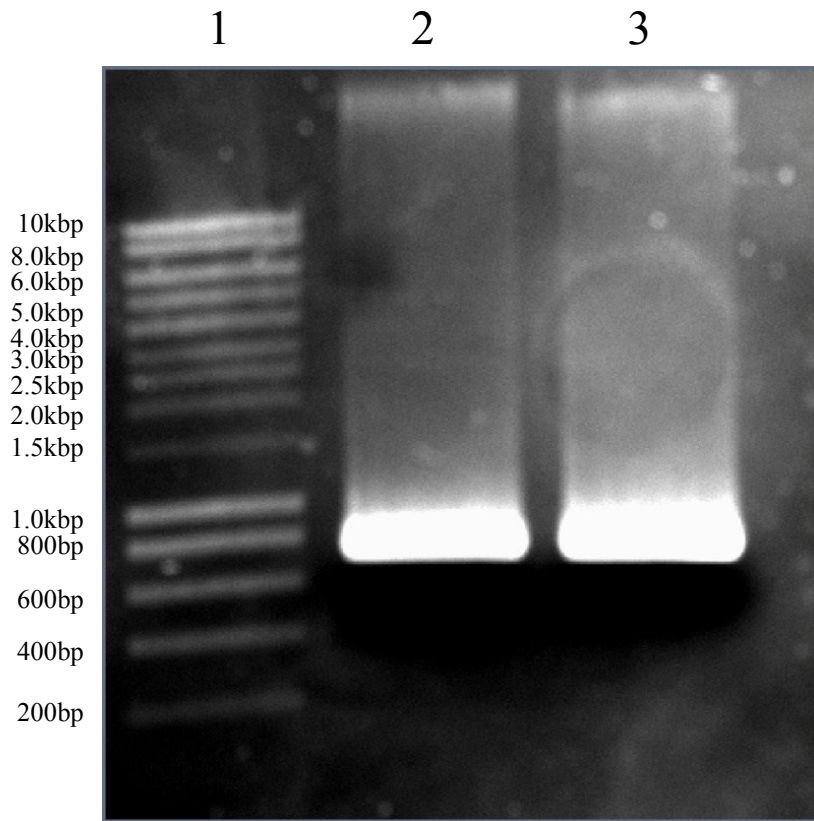


Figure 4.4: PCR amplification of *seq0849*.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standard; lane 2, PCR product from amplification of *seq0849* from *S. equi* 4047 genomic DNA, Lane 3; PCR product technical replicate.

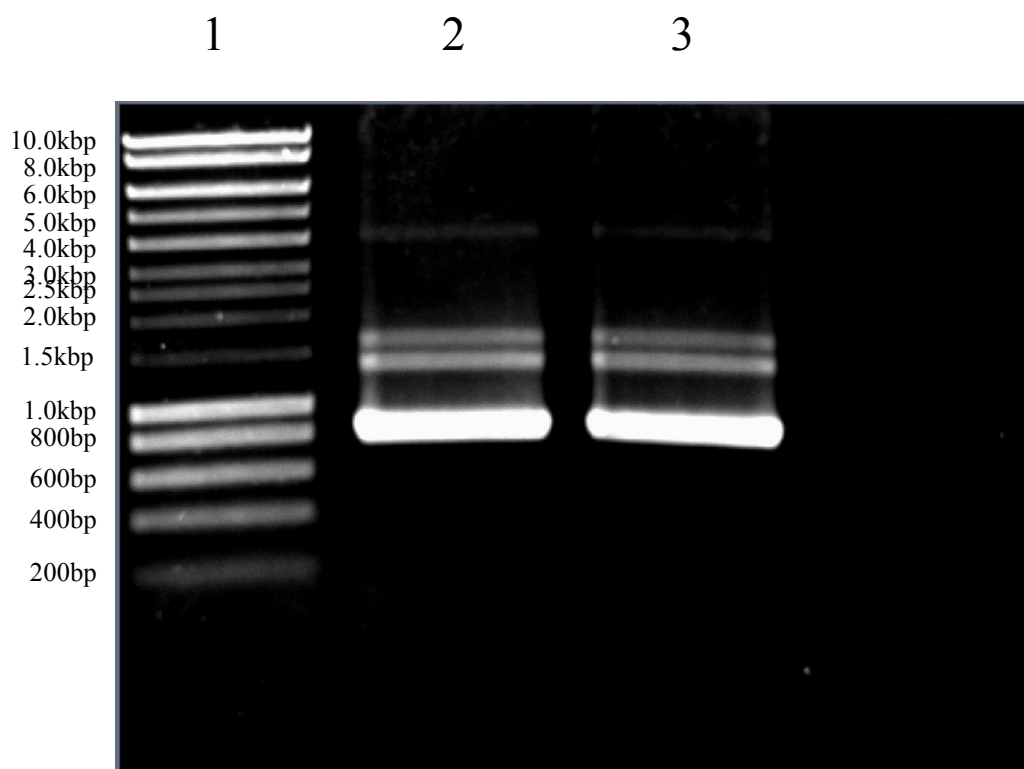


Figure 4.5: PCR amplification of *seq2115*.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 Kbp size standard; lane 2, PCR product from amplification of *seq2155* from *S. equi* 4047 genomic DNA; lane 3, PCR product technical replicant.

4.4. Cloning of *seq0849* and *seq2155*

The amplified DNA products were initially cloned into pCR[®]-Blunt vector, yielding plasmids pb0849 and pb2155. These plasmids were subsequently digested with appropriate restriction endonucleases to verify product integration. Cloned target gene and pCR-Blunt plasmid DNA was then digested with appropriate endonucleases, purified and cloned into similarly restriction digested and purified pET 28a vector. This generated plasmids p0849 and p2155 that encoded the proteins of interest with a His₆-tag at the N-terminus of the mature protein.

Confirmation of the correct plasmid construction was undertaken by screening of plasmids extracted from candidate positive clones with appropriate restriction endonucleases. Following digestion, a successful clone was indicated by the visualisation of a product consisting of the expected target gene size and the remaining restriction endonuclease cut pET 28a vector, 5.4 kbp in size (Figures 4.6).

Further confirmation was undertaken by the gene sequencing of the plasmid from the T₇ promoter region. The sequences obtained were checked for fidelity and aligned to their respective nucleotide sequences in the *S. equi* 4047 genome (See appendix IV). Dr Meng Zhang is thanked for technical assistance with some of the steps in this cloning procedure.

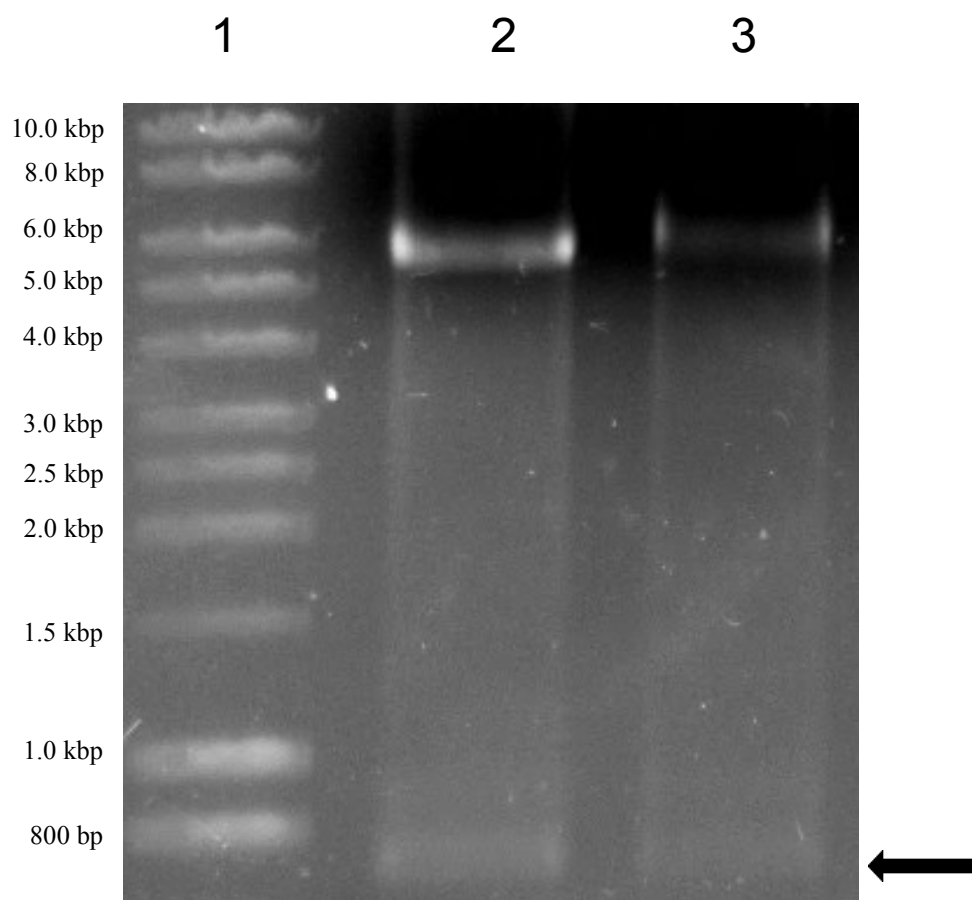


Figure 4.6: Restriction endonuclease digestion of p0849 and p2155

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 $\mu\text{g/ml}$). Lanes were loaded as follows: Lane 1, 10 kbp size. Lane 2, p0849 post treatment with *NdeI* and *XhoI* restriction endonucleases. Lane 3, p2155 post treatment with *NdeI* and *XhoI*. Released *seq0849* and released *seq2155* product (expected sizes 715bp and 779bp respectively) are indicated with an arrow.

4.5. Protein expression of SEQ0849 and SEQ2155

Plasmids p0849 and p2155, encoding SlaA and SlaB respectively, were chemically transformed into a range of *E. coli* expression strains (see Table 2.1) and grown at 37 °C to mid-exponential phase. Initially when plasmid, p0849 was transformed into a variety of *E. coli* expression strains, expression of the recombinant proteins was induced by the addition of 1 mM IPTG (final concentration) and the cultures were incubated at a variety of temperatures. A temperature of 20°C was initially chosen as a reduced temperature is a known technique to limit the *in vivo* aggregation of recombinant proteins (Schein 1989).

Whole cell protein extractions visualised by SDS-PAGE yielded poorly expressed target protein (Figure 4.7). The cultivation temperature was therefore increased to 37°C as it has been shown that increased temperature of cultivation can result in an increased level of replication, transcription and translation (Shaw *et al* 1967). It was however shown that expression at 37°C produced a notably reduced level of expressed soluble protein compared to cultures incubated at 20°C (Figure 4.8).

As well as the two common *E. coli* expression strains (BL21 and B834), two strains were chosen specifically for this expression, Tuner™ and C43 (DE3). The expression strain *E.coli* Tuner™ is useful since this strain enables the adjustment of protein expression by adjusting the concentration of IPTG. A reduced level of expression has been shown to enhance the solubility and activity of difficult target proteins (Novagen 2003). The expression strain C43 (DE3) can significantly increase soluble protein expression of some difficult recombinant proteins that are unable to express in the parent strain BL21 (DE3)

(Miroux, 1996). The strain C43 (DE3) has also been demonstrated to increase the stability of plasmids encoding toxic proteins (Dumon-Seignover *et al* 2004).

A further area of investigation was the use of glucose to control basal expression in the pET system. It has been demonstrated that the addition of 0.1% glucose to cultures in LB medium can increase the expression of a target protein (Novy and Morris 2001).

In summary, SDS-PAGE analysis indicated that expression of SlaA was found to be most favourable when expressed in the *E. coli* expression strain Tuner™, induced with 0.5 mM IPTG (final concentration) followed by the addition of 0.1% w/v glucose, agitated at 100rpm at 20°C overnight (Figure 4.9).

Expression of SlaB recombinant protein from plasmid p2155 was also found to be most favourable under conditions similar to that of the SlaA recombinant protein (Figure 4.10). The molecular weight of the mature SlaA and SlaB proteins including the His₆ N-terminal tag was estimated to be 20.6 kDa using the Expasy ProtParam tool (<http://expasy.org/tools/protparam.html>).

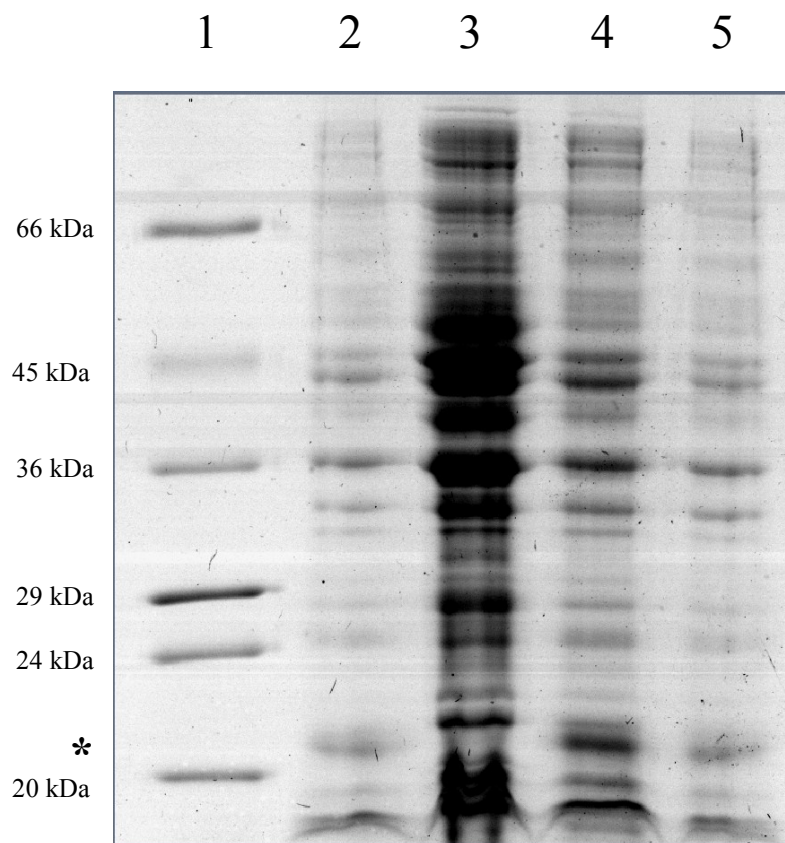


Figure 4.7: 17% SDS PAGE of SlaA expression trials.

Protein expression was induced by 1 mM IPTG and growth at 20°C. The expected position of the SlaA protein is marked with an asterisk. Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 20kDa); lane 2, *E.coli* Tuner (DE3) p0849; lane 3, *E. coli* C43 (DE3) p0849; lane 4, *E. coli* B834 (DE3) p0849; lane 5, *E. coli* BL21 (DE3) p0849.

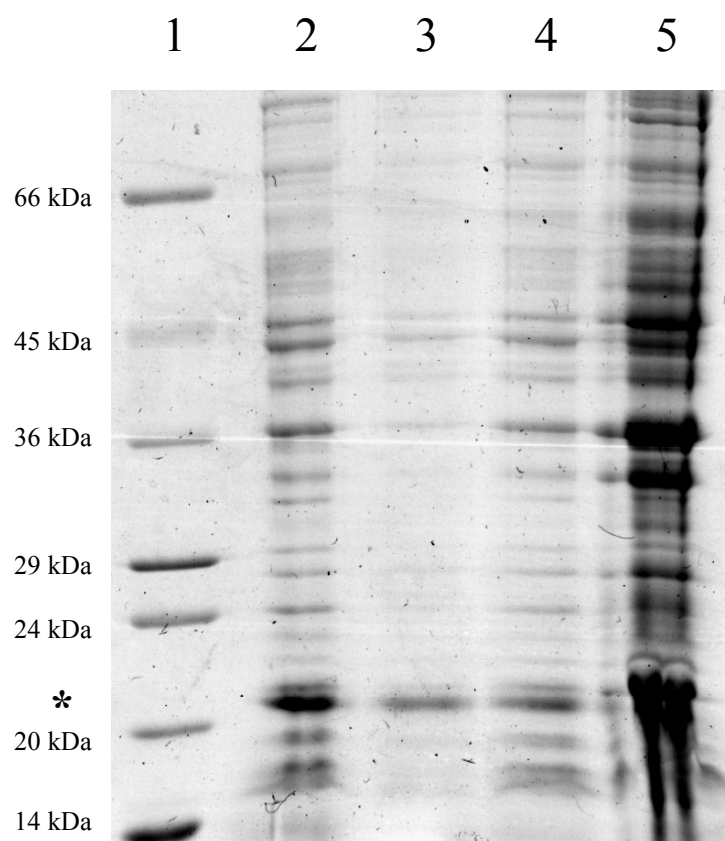


Figure 4.8: 17% SDS PAGE of SlaA expression trials.

Protein expression was induced by 1 mM IPTG. The expected position of the SlaA protein is marked with an asterisk. Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 20 kDa); lane 2, *E. coli* BL21 (DE3) p0849 induced at 20 °C; lane 3, *E. coli* BL21 (DE3) p0849 induced at 20 °C overnight; lane 4, *E. coli* BL21 (DE3) p0849 induced at 37 °C; lane 5, *E. coli* BL21 (DE3) p0849 induced at 37 °C overnight.

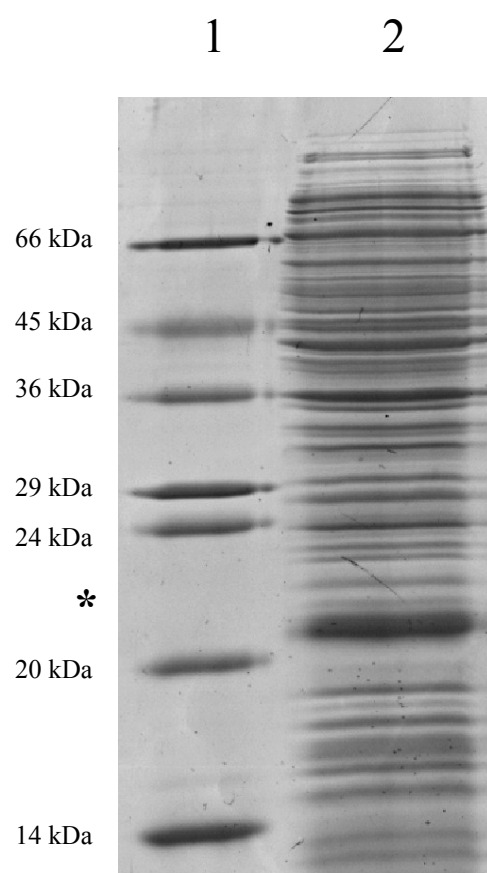


Figure 4.9: 17% SDS PAGE of SlaA expression.

Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 20, 14 kDa); lane 2, *E. coli* Tuner® p0849 induced at 20 °C overnight with the addition of 0.1% w/v D-Glucose. Induced SlaA protein is marked with an asterisk.

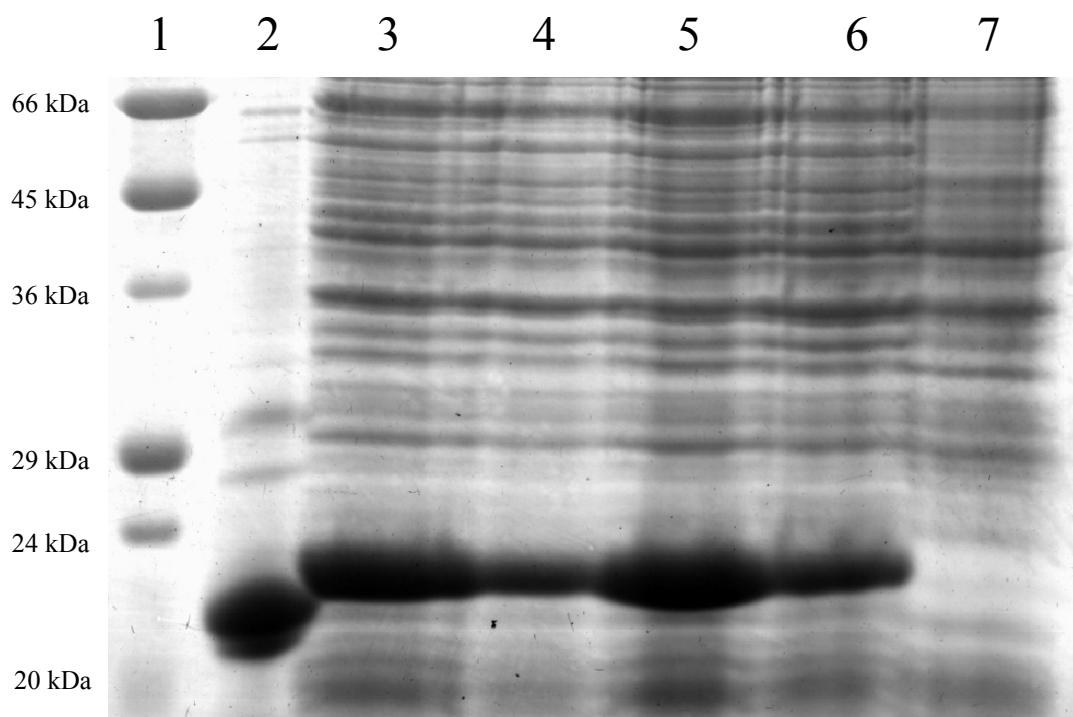


Figure 4.10: 17% SDS PAGE of SlaB expression trials.

SlaB expression was induced by the addition of 0.1% w/v D-Glucose. Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 20 kDa); lane 2, Control (AKTA Purified SlaA); lane 3, *E. coli* BL21 (DE3) expressed at 20°C overnight; lane 4, *E. coli* BL21 (DE3) expressed at 37°C; lane 5, *E. coli* Tuner (DE3) expressed at 20°C overnight; lane 6, *E. coli* Tuner (DE3) expressed at 37°C overnight; lane 7, control (*E. coli* BL21 with p2155 plasmid not present).

4.6. Protein purification of SlaA and SlaB

Both His₆-tagged proteins were expressed in 4 litres of LB media and purified from cell free extracts by (IMAC) using a chelating Sepharose column and an automated gradient elution technique (Section 2.13.4). Confirmation of proteins present in different fractions was indicated by U.V absorption; fractions were checked for purity by SDS-PAGE (Figures 4.11 and 4.12) and concentrated. Proteins were then further purified by FPLC, monitored by U.V absorption and checked for purity by SDS-PAGE, concentrated and quantified (Figures 4.13 and 4.14). Quantities of purified SlaA obtained were ~3.75 mg/L. It was noted that quantities of purified SlaB obtained from recombinant protein expression was significantly less than comparable expression quantities of SlaA (~0.125 mg/L).

The purified SlaB product was also noted to have a lower stability compared with similarly obtained purified SlaA, showing visible degradation on SDS-PAGE analysis during a shorter time-frame. Additionally a reduced level of purification was observed showing residual *E. coli* protein present at a higher molecular weight than the target protein on SDS-PAGE analysis (Figure 4.15). A serine specific protease inhibitor was added during recombinant expression of SlaB but was found to have little effect on the quantity or quality of recombinant protein obtained (data not shown)

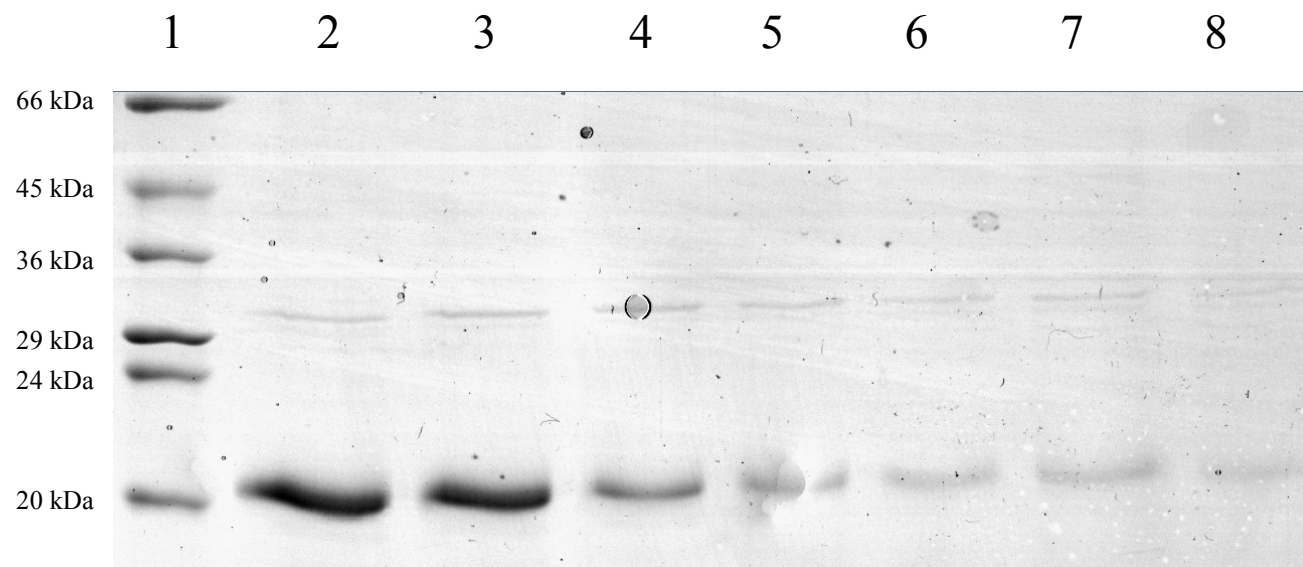


Figure 4.11: 17% SDS PAGE of IMAC purified fractions of SlaA.

Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 24, 20kDa); lanes 2-8, individual fractions of purified SlaA (Fractions 49, 51, 53, 55, 57, 59, 61).

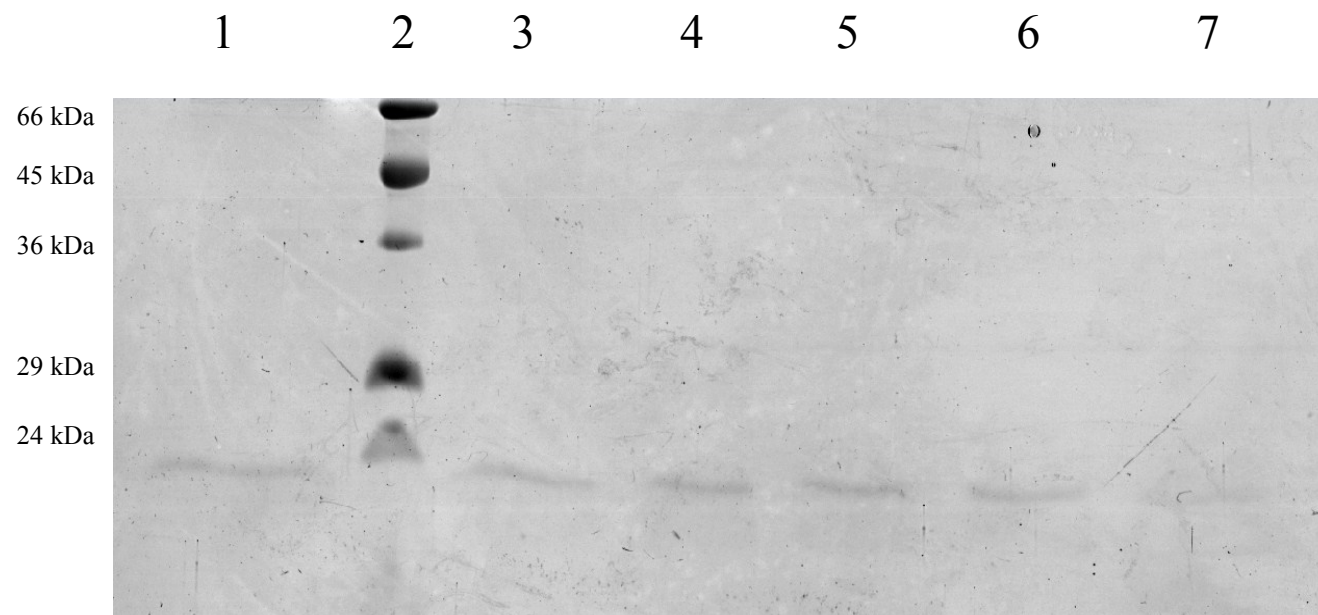


Figure 4.12: 17% SDS PAGE of IMAC purified fractions of SlaB.

Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 24, 20 kDa); lanes 2-7, individual fractions of purified SlaB (Fractions 55, 57, 59, 61, 63, 65).

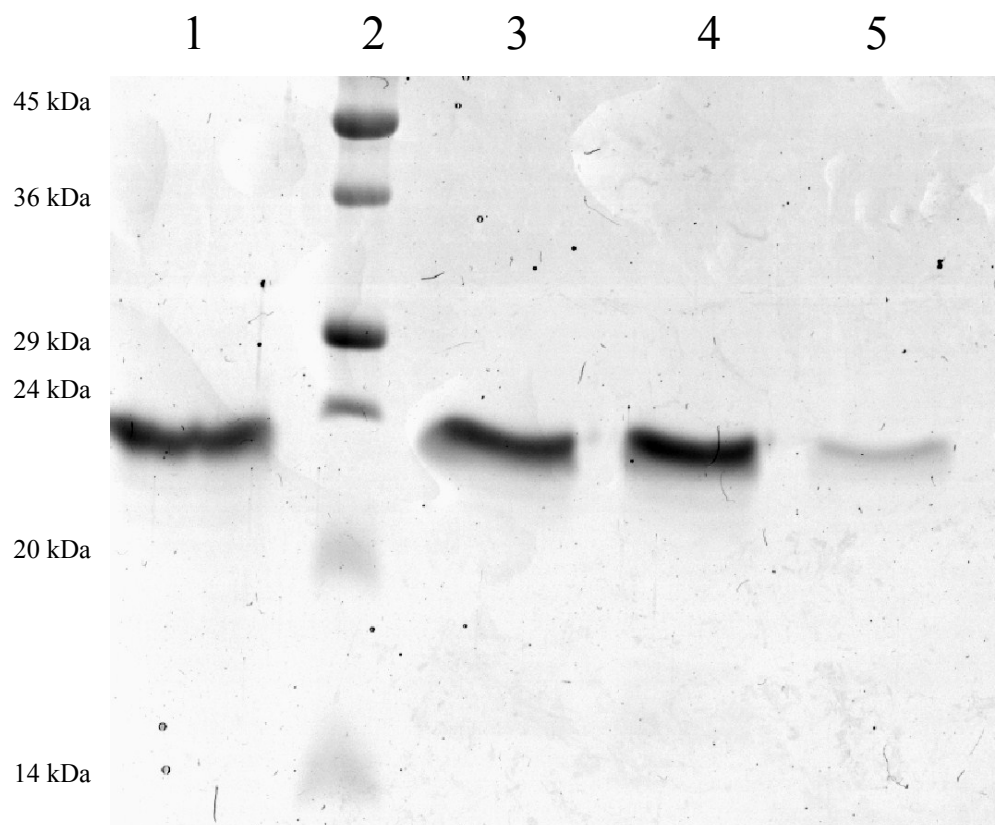


Figure 4.13: 17% SDS PAGE of FPLC purified fractions of SlaA.

Lanes were loaded as follows: Lane 1 purified SlaA (fraction 23); lane 2, molecular weight markers (45, 36, 29, 24, 20, 14 kDa); lanes 3-5, individual fractions containing purified SlaA (fractions 24, 25, 26).

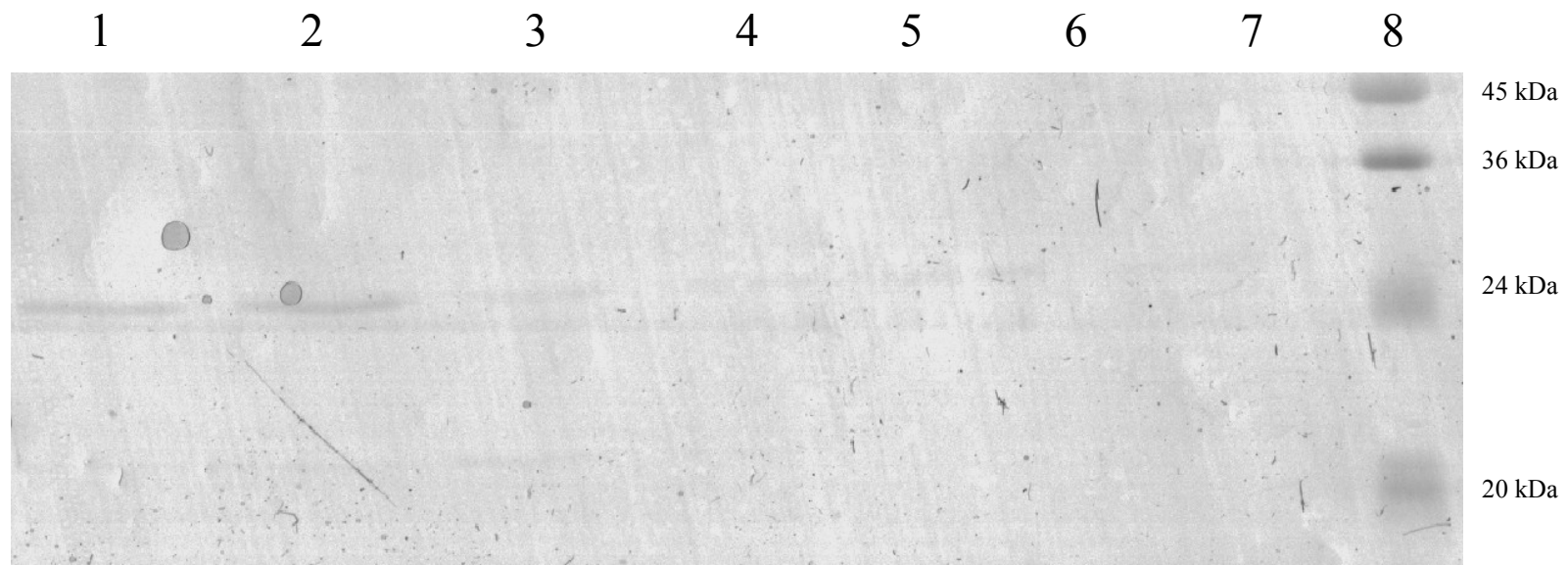


Figure 4.14: 17% SDS PAGE of FPLC purified fractions of SlaB.

Lanes were loaded as follows: Lanes 1-7, individual fractions containing purified SlaB (23, 24, 25, 26, 27, 28, 29 kDa); lane 8, molecular weight markers (45, 36, 24, 24, 20 kDa).

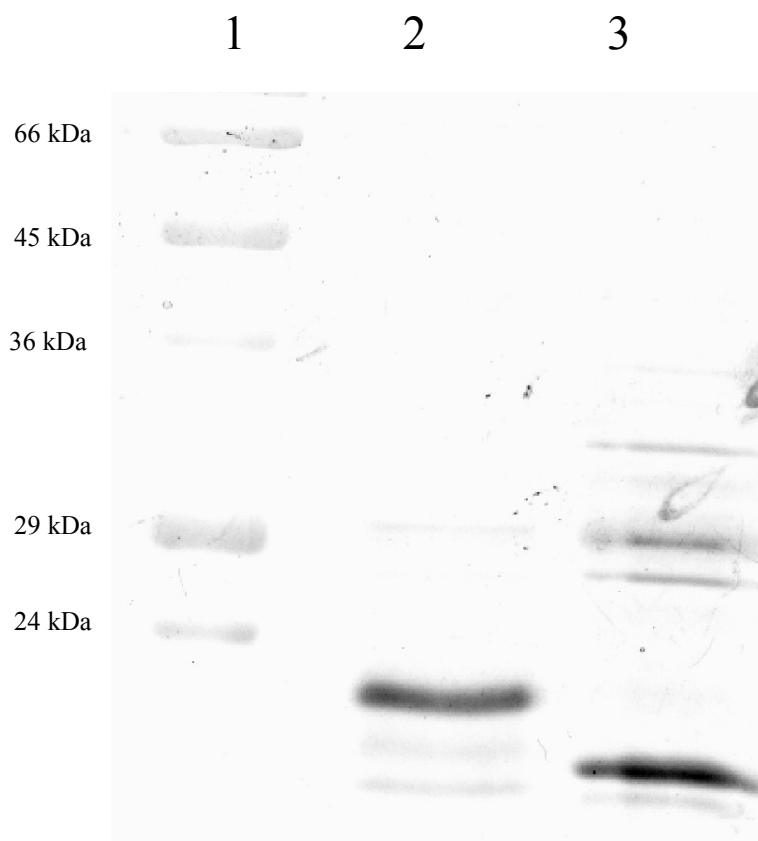


Figure 4.15: 17% SDS PAGE of the purified SlaA and SlaB products obtained

Lanes were loaded as follows: Lane 1, molecular weight markers (66, 45, 36, 29, 24, 24, 20kDa); lane 2, purified SlaA; lane 3, purified SlaB. 2 μ g of purified protein was added to each sample lane.

4.7. Confirmation of His₆-Tagged SlaA and SlaB by western blotting

The presence of His₆-Tagged SlaA and SlaB in the protein samples obtained was screened by Western immunoblotting, using a commercial polyclonal antiserum raised against the His₆-tag. The antiserum was shown to cross-react with a ~20kDa proteins in each preparation consistent to the predicted size of SlaA and SlaB. Cross-reaction was especially strong against SlaA, with a less intense reaction shown against SlaB. Notably a reaction was also detected above and below the expected band size of ~20kDa against SlaA. The reaction above is likely to represent protein aggregation products and the reaction below is likely to represent protein degradation products (Figure 4.16).

4.8. Detection of antibodies to SlaA and SlaB in equine convalescent serum

The presence of antibodies to SlaA and SlaB in equine convalescent serum was investigated by Western immunoblotting, using equine convalescent serum acquired from three ponies, pre and post infection with *S. equi*. Strong antibody cross-reactivity with SlaA was shown in one of the serum pony samples acquired post-infection and a very weak reaction shown against SlaB. No cross reactivity was detected in serum acquired pre-infection against either SlaA or SlaB (Figure 4.17). Note that two different molecular weight markers were used in each separate panel. The left hand panel (lanes 1-3) used a pre-stained molecular weight marker whereby the right hand panel used a marker that was only visible using reversible staining with Ponceau reagent (marked in pencil prior to western blot development).

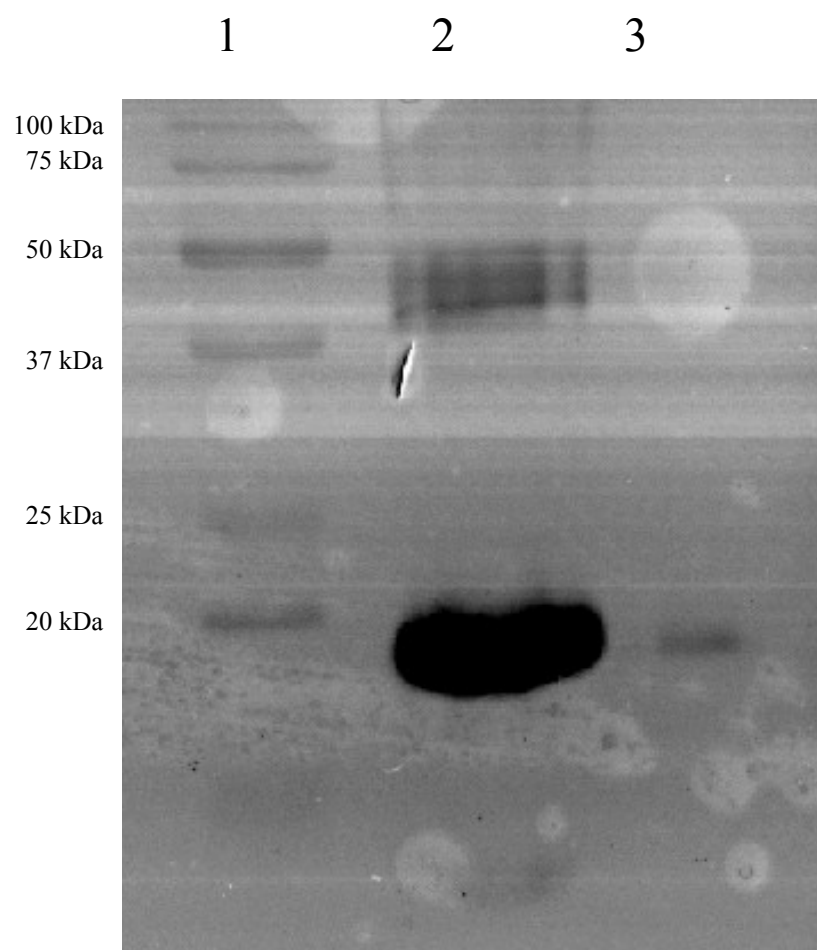


Figure 4.16: His₆-tag Western blot analysis of SlaA and SlaB.

Lanes were loaded as follows: Lane 1, molecular weight markers (20, 25, 37, 50, 75, 100 kDa); lane 2, SlaA. The purified protein amount added was 2 μ g.

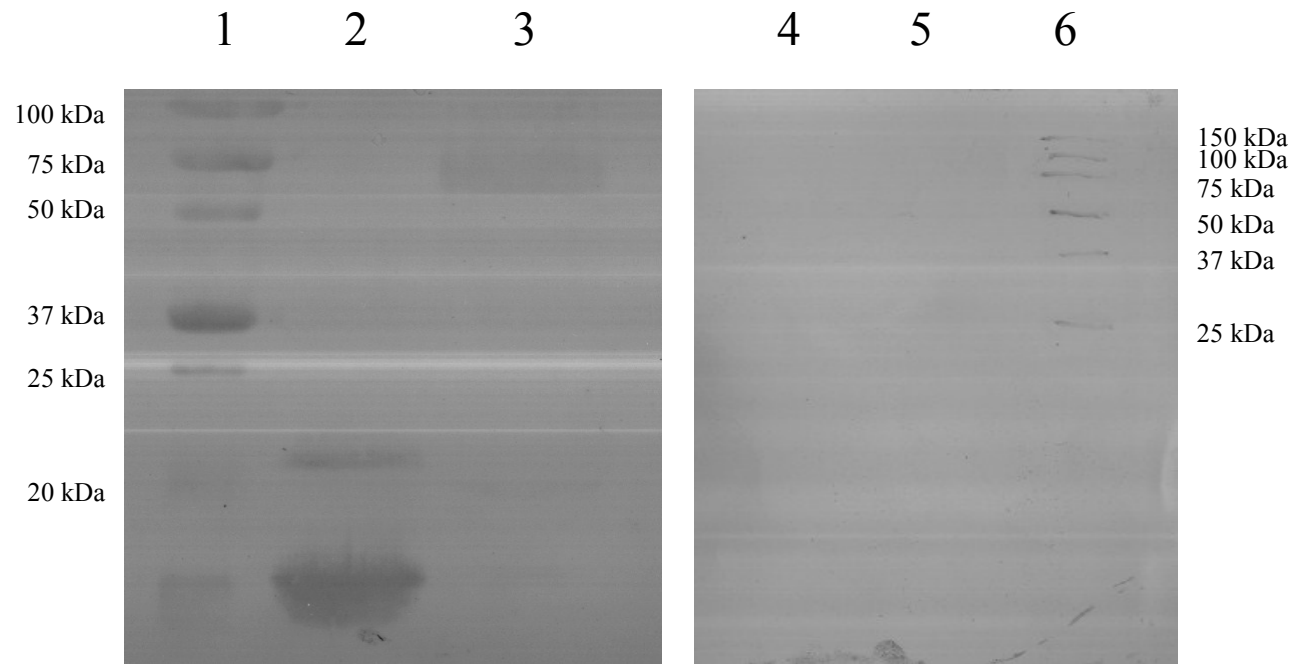


Figure 4.17: Western blot analysis of purified SlaA and SlaB against equine convalescent serum (pre and post infection).

Lanes were loaded as follows: Lane 1 and 6: molecular weight markers (20, 25, 37, 50, 75, 100 kDa); lane 2, SlaA; lane 3, SlaB; lane 4, SlaA; lane 5, SlaB. Amount of protein added to each well was 2 μ g. Left hand pannel (lanes 1-3), blot probed with post-infection serum (from Pony 5080). Right hand panel (lanes 4-6), blot probed with matched pre-infection serum.

4.9. Enzymatic assay of SlaA and SlaB activity

4.9.1. **Sigma method**

For investigation of the activity of SlaA and SlaB against L- α -phosphatidylcholine, the initial assay method recommended by Sigma-Aldrich[®] was undertaken (see section 2.14.2.1). A positive control sPLA₂ enzyme (source – bee venom) was employed to test the robustness of the chosen assay. Initially a calibration curve was calculated using L- α -phosphatidylcholine as the standard. Subsequently positive control sPLA₂ and purified SlaA protein was tested, due to a superior quantity available in comparison to SlaB. The total amount of both enzymes added to each reaction mixture was 5 μ g at 37°C and absorbance was read at 570 nm. Although visual colorimetric changes were observed in the initial series of assays for both enzymes, the data obtained did not reflect true Michaelis-Menten kinetics. Furthermore poor enzyme stability was observed and a large degree of variation amongst replicates was observed. This assay method was therefore deemed inappropriate for enzyme activity quantification (data not shown).

4.9.2. **Price method**

A second assay method detailed by Price (2007) was investigated (see section 2.14.2.2), as a viable cost-effective alternative. As the assay detects phospholipase activity from the release of protons, initial evaluation of the assay was conducted using measured amounts of HCl in

place of the enzyme. The range of the assay when HCl was used as an evaluator reagent was shown to be linear but overall stability of the reaction mixture and reproducibility of the dataset was poor (Figure 4.18).

Regarding sPLA₂ enzymatic activity, initial assays were undertaken using SlaA to assess the optimum viable enzyme concentration suitable for further investigations with a standard final concentration of 5 mM phosphatidylcholine substrate. The initial assay assessment indicated that an amount of 2 µg was most appropriate although again general poor reproducibility of results was observed (Figure 4.19). Finally a further investigation of the effect of substrate concentration was undertaken with this assay method. Data indicated that true Michaelis-Menton kinetics for enzyme activity was not observed for this assay when undertaken with SlaA (Figure 4.20). The Assay was also performed with 0.2 µg of SlaB, although visually a colorimetric reaction was observed, the absorbance change recorded was not significant (data not shown).

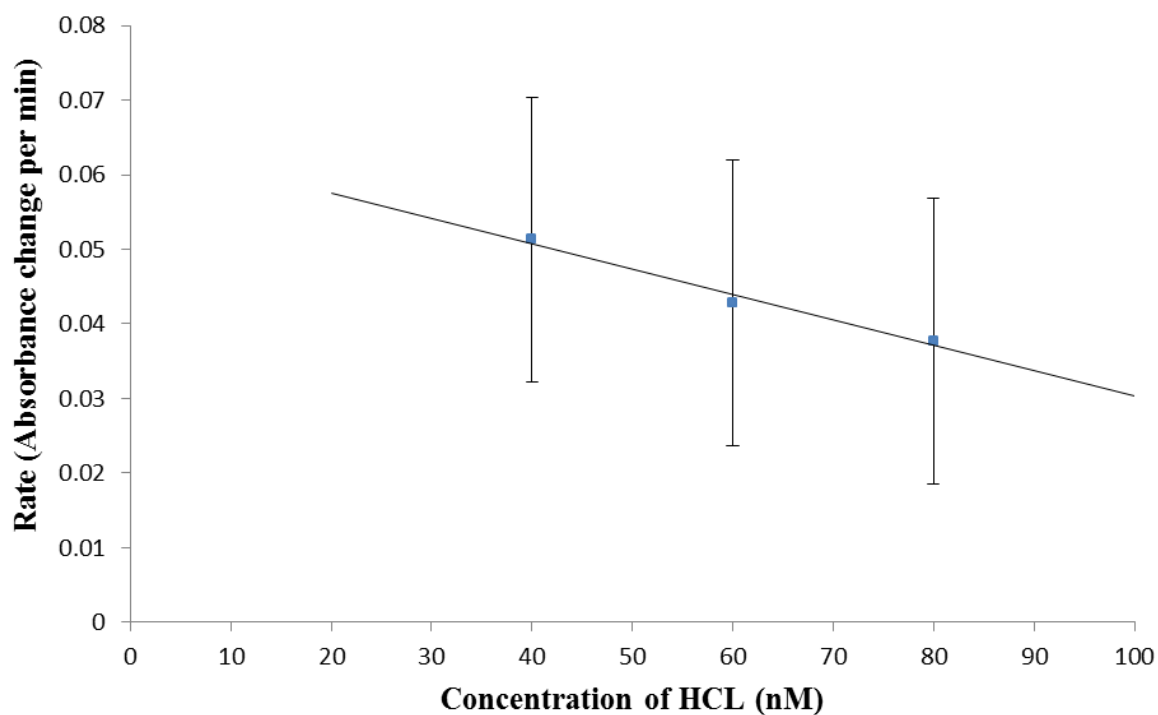


Figure 4.18: Enzyme assay by the method of Price (2007).

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx 10 minutes with varying concentrations of HCL present and a standard concentration of 5mM of substrate. Error bars represent the standard deviation from the mean.

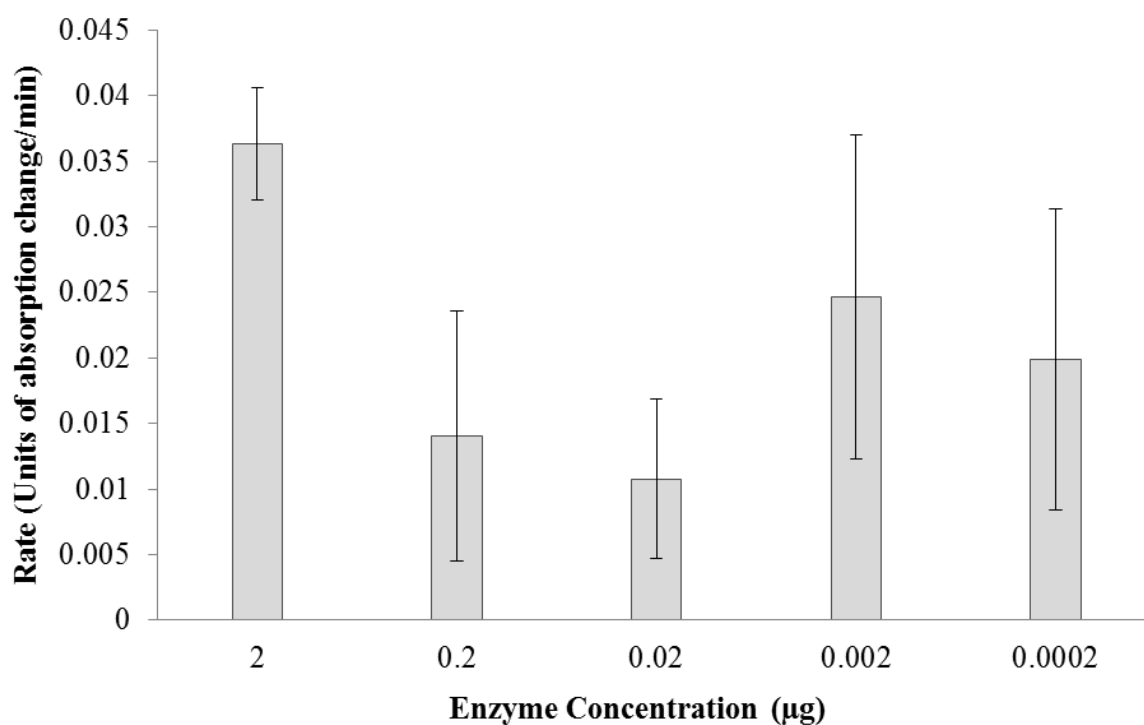


Figure 4.19: Enzyme assay by the method of Price (2007).

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx 10 minutes with varying amounts of SlaA enzyme and a standard concentration of 5 mM of substrate. Error bars represent the standard deviation from the mean.

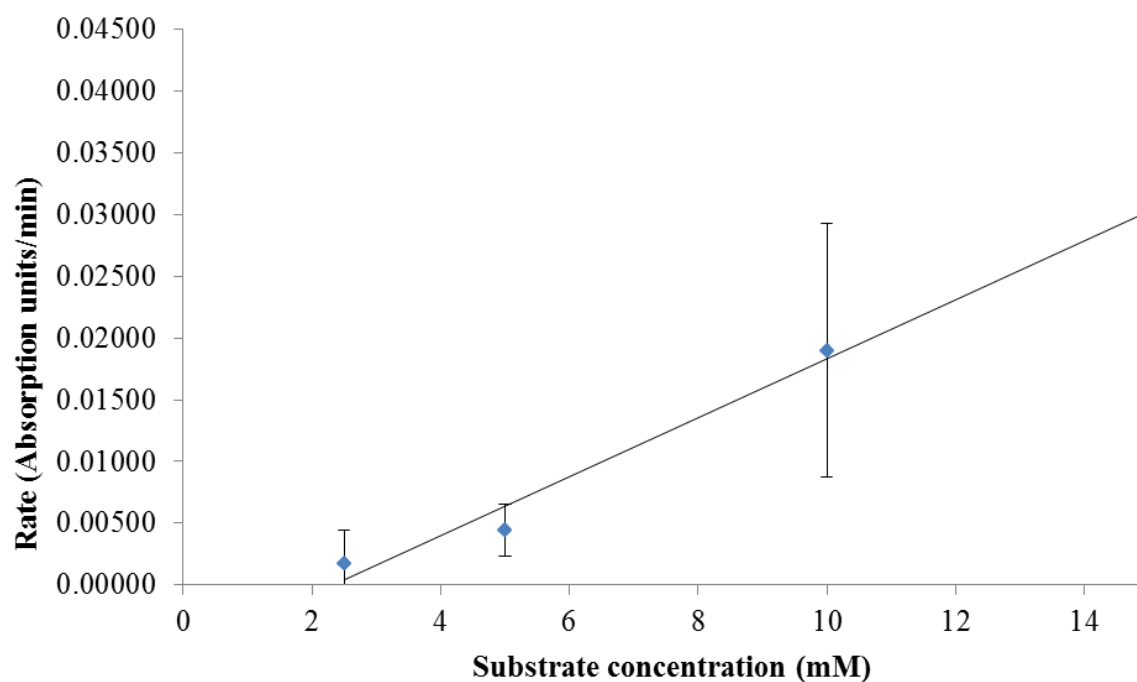


Figure 4.20: sPLA₂ enzyme assay by the method of Price (2007).

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx 10 minutes with 2.0 µg of active SlaA. Error bars represent the standard deviation from the mean.

4.9.3. Chromogenic substrate release assay

A further assay was investigated that was based on the work of Reynolds (1992), using a commercial kit obtained from Cayman Chemical Company and detailed in section 2.14.2.3. Initially the assay was assessed for linearity and accuracy using 1 μg of a commercially available enzyme control and a substrate concentration of 1.66 mM (PLA₂ obtained from Bee venom). The reaction was allowed to proceed at room temperature with the absorption measured at 414 nm for 10 minutes and demonstrated to be linear (Figure 4.21). The specific activity of Bee venom PLA₂ against diheptanoyl thio-phosphatidylcholine (DHT-PC) was calculated to be $1.743 \pm 0.19 \mu\text{mol}/\text{min}/\text{mg}$.

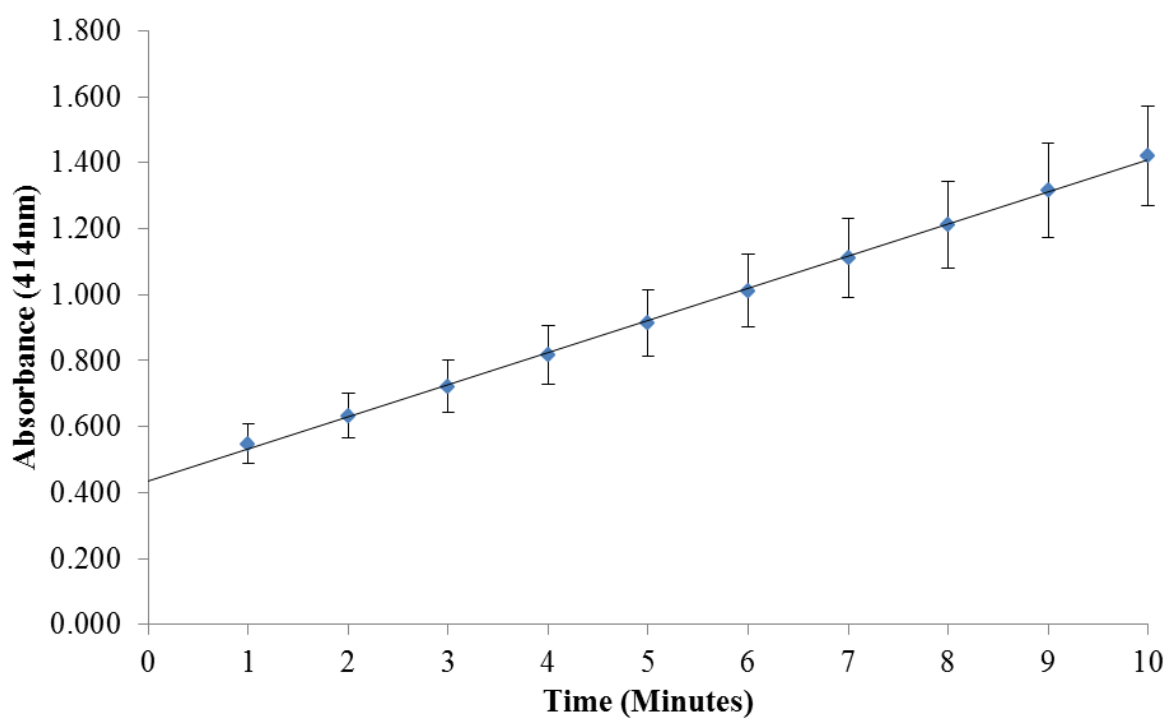


Figure 4.21: Standard rate of sPLA₂ activity over against time.

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx 10 minutes with 1.0 μg of active sPLA₂ (source- bee). Error bars represent the standard deviation from the mean.

4.9.3.1. Michaelis-Menten parameters for SlaA

For the investigation of the activity of SlaA against DHT-PC, the enzyme was assayed for an increase in absorbance at 414 nm. A total amount of 10 µg of enzyme (determined empirically) was added to each reaction to assay SlaA and a total amount of 1 µg of enzyme was added to each reaction to assay SlaB.

For kinetic analysis, 4 substrate concentrations (determined empirically) were chosen and reactions performed in triplicate with changes in absorbance observed over a total of ten minutes (Figure 4.22). The calculated K_m for SlaA was determined to be 14.40 ± 7.866 mM and demonstrated on a Lineweaver-Burk plot (Figure 4.23). Specific activity was determined to be $5.06 \times 10^{-2} \pm 3.01 \times 10^{-3}$ µmol/min/mg.

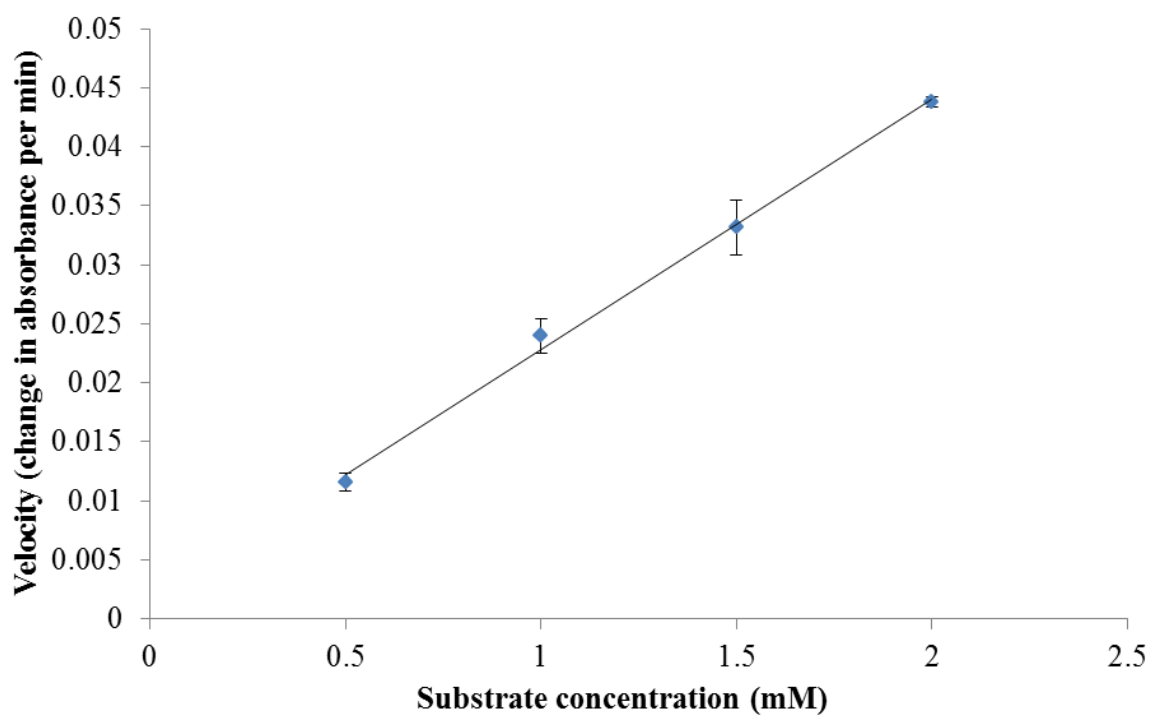


Figure 4.22: Rate of SlaA activity against substrate concentration.

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for 10 minutes with 10.0 μg of active SlaA. Error bars represent the standard deviation from the mean.

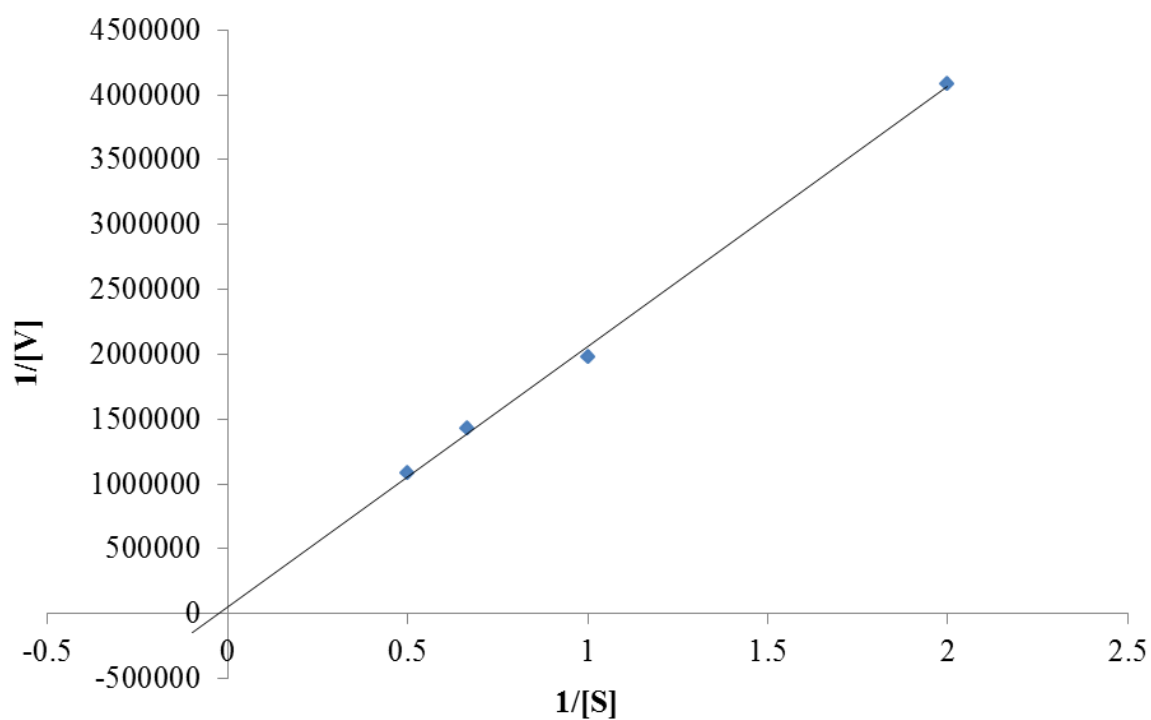


Figure 4.23: Lineweaver-Burk plot for SlaA against DHT-PC.

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for 10 minutes with 10.0 μg of active SlaA.

4.9.3.2. Determination of temperature optimum

An investigation was undertaken to assess the optimal temperature of enzyme activity for SlaA. A total of 10 µg of enzyme and 1 mM of DHT-PC substrate were used in each reaction and the reaction was undertaken for a period of 10 minutes.

As shown in figure 4.24, SlaA demonstrated the highest rate of activity against DHT-PC at 45°C, ($6.1 \times 10^{-2} \pm 2.3 \times 10^{-3}$ µmol/min/mg). Significant standard deviation of the means was demonstrated at 50 °C ($\pm 5.6 \times 10^{-3}$ µmol/min/mg) perhaps suggesting enzyme instability or degradation at the higher temperatures.

4.9.3.3. Determination of divalent ion requirement

When SlaA was tested with a range of divalent cations, it was observed that the enzyme had a specific requirement for Ca^{2+} ions. A final concentration for each divalent ion added was 10mM, which was noted as the standard amount in reference sPLA₂ assays.

The data for the effect of divalent cations on activity is presented in figure 4.25. No activity was recorded when activity was investigated with either Cu^{2+} or Mn^{2+} , slight activity was recorded when Co^{2+} replaced the standard Ca^{2+} although this varied considerably amongst replicates and not deemed significant.

The replacement of Ca^{2+} with Mg^{2+} revealed a quantifiable level of specific activity ($6.8 \times 10^{-3} \mu\text{mol}/\text{min}/\text{mg}$). Comparisons of the control and replacement of Ca^{2+} with Mg^{2+} demonstrated that the rate of enzymatic activity was reduced by approximately 88.2%.

The effect of increased concentration of Ca^{2+} was also demonstrated and showed a significant measurable increase when the concentration of available divalent cation's was increased 4-fold to a final concentration of 40 mM. An increase of 76.3% in activity was observed from an enzyme activity of $4.34 \times 10^{-2} \pm 2.4 \times 10^{-3} \mu\text{mol}/\text{min}/\text{mg}$ to $0.181 \pm 5.3 \times 10^{-3} \mu\text{mol}/\text{min}/\text{mg}$ (Figure 4.26).

A reduction of Ca^{2+} concentration in the reaction (below the standard concentration of 10mM) was found to have minimal effect on enzyme activity. The rate of activity was demonstrated to have little variation between the standard concentration of 10 mM and a reduced concentration of 2.5 mM. Unexpectedly a small amount of activity was detected in the absence of any Ca^{2+} ($3.23 \times 10^{-3} \pm 4.9 \times 10^{-4} \mu\text{mol}/\text{min}/\text{mg}$).

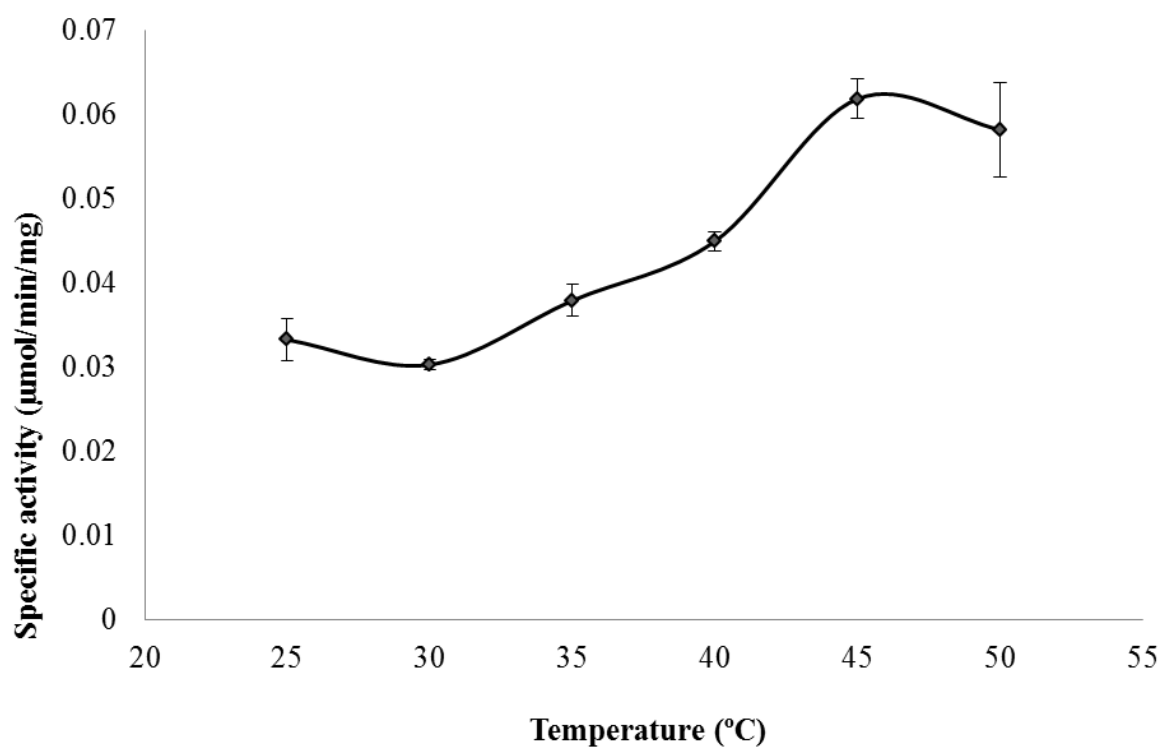


Figure 4.24: The effect of temperature on the rate of SlaA enzyme activity against DHT-PC

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction had a final substrate concentration of 0.5 mM and 10.0 μg of active SlaA. Error bars represent the standard deviation from the mean.

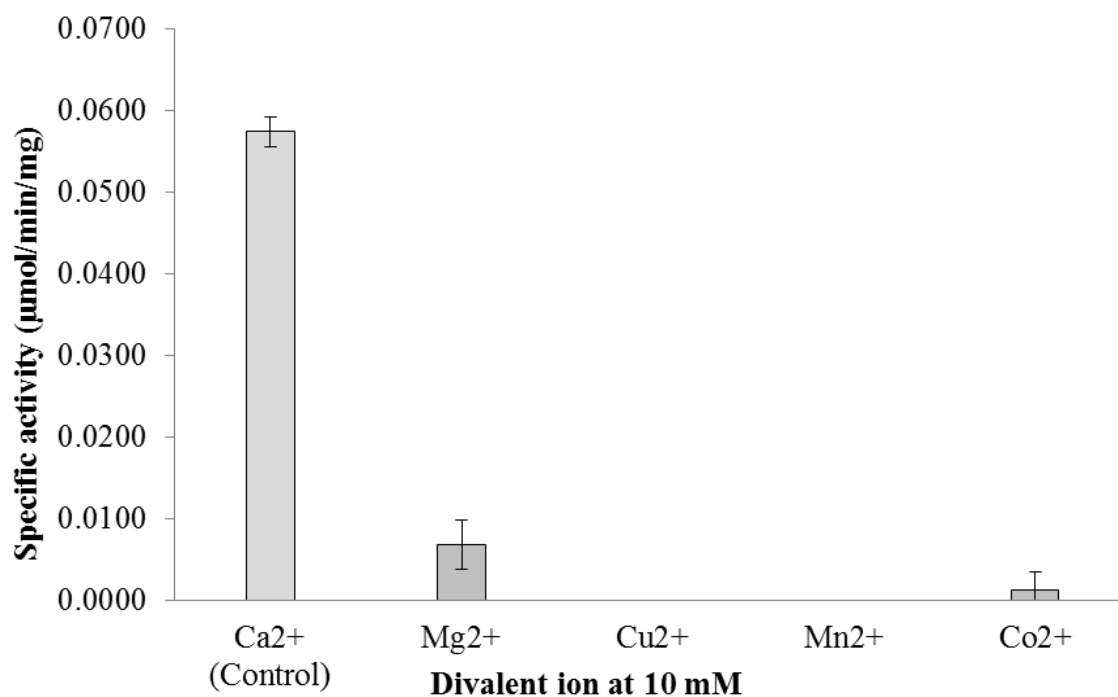


Figure 4.25: The effect of divalent ion replacement on the rate of SlaA enzyme activity.

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction had a final DHT-PC substrate concentration of 0.5 mM and 10.0 μg of active SlaA. The control contained 10 mM Ca²⁺ ions. Error bars represent the standard deviation from the mean.

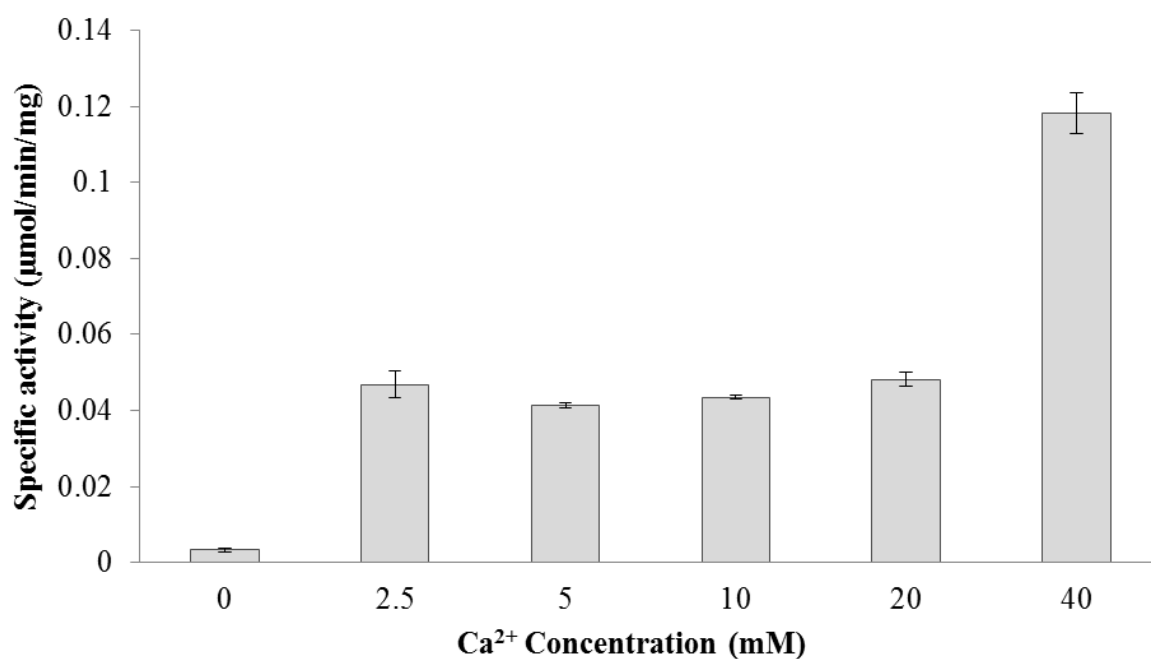


Figure 4.26: The effect of calcium ion concentration on the rate of SlaA enzyme activity.

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction had a final DHT-PC substrate concentration of 0.5 mM and 10.0 μg of active SlaA. Error bars represent the standard deviation from the mean.

4.9.3.4. Determination of SlaA and SlaB activity against arachidonyl thio-PC

Cayman Chemical Co. provides a kit for sPLA₂ assay with arachidonyl thio-PC (AT-PC) as an alternative substrate. Enzymatic activity of control Bee sPLA₂ was initially tested against AT-PC at a final substrate concentration of 1 mM and an enzyme quantity of 1 µg. The enzymatic rate of activity of 1 µg of Bee sPLA₂ against AT-PC, compared to the standard DHT-PC substrate was demonstrated to be reduced by 87.4% ($0.243 \pm 7.4 \times 10^{-3}$ and $1.887 \pm 9.2 \times 10^{-3}$ µmol/min/mg respectively; figure 4.27). The enzymatic rate of activity of SlaA and SlaB was assayed against AT-PC with 10.0 µg of SlaA and 1 µg of SlaB at 37°C. It was however demonstrated that neither SlaA nor SlaB possessed any significant activity against AT-PC.

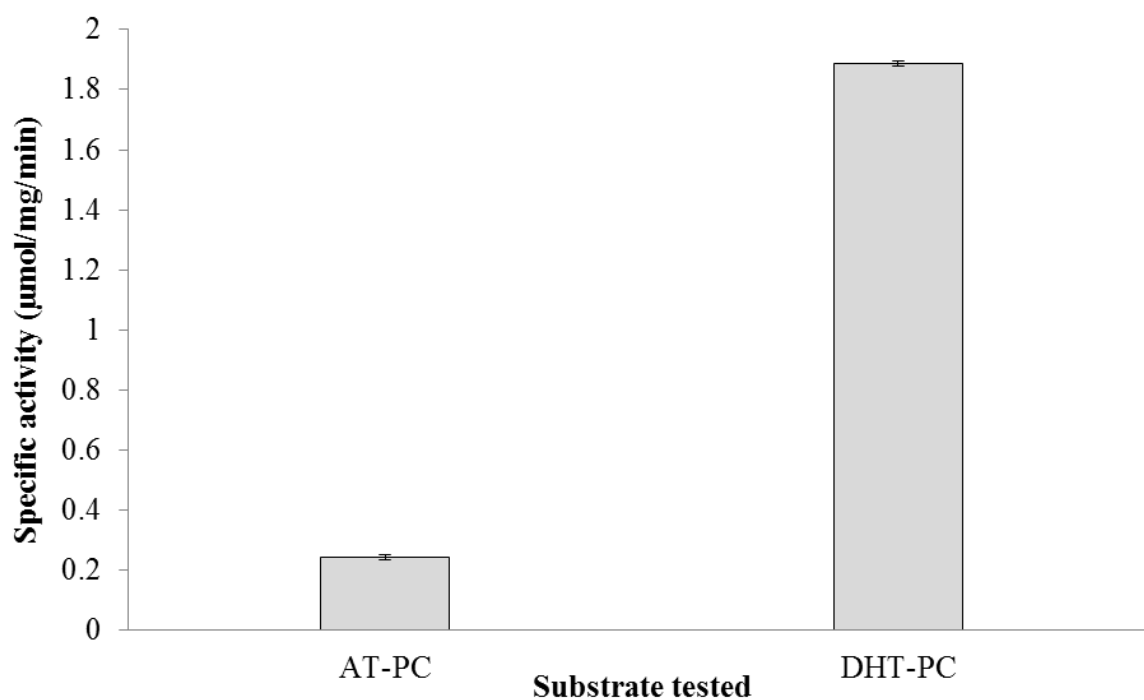


Figure 4.27: sPLA₂ specific enzyme activity of positive control phospholipases against different lipid substrates.

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction had a final substrate concentration of 1.0 mM and 10.0 μg of active bee sPLA₂. Error bars represent the standard deviation from the mean.

4.10. Discussion and future work

Bioinformatical analysis showed that *S. equi* encodes two sPLA₂ enzymes, designated *slaA* and *slaB*. Furthermore it was shown that the closely related *S. zooepidemicus* encodes only an orthologue of *slaB*, with that of *slaA* absent. It would be reasonable to speculate that an ancestral organism of *S. equi* and *S. zooepidemicus* possessed both *slaA* and *slaB* prior to the sub-species divergence event. Logically, *S. zooepidemicus* then probably lost *slaB* during this process; this is consistent with *S. zooepidemicus* possessing a smaller genome with fewer predicted CDSs (Holden *et al.* 2009). The presence of SpSlaA is thus likely to have occurred via acquisition of SlaA from *S. equi* since a SpSlaA strain of *S. pyogenes* has yet to be identified that predates the emergence of SlaA in *S. equi* (Sitkiewicz *et al.* 2007). Although an alternative scenario could indeed be that a *S. equi* ancestral stain possessed only SlaB and SlaA was a result of a duplication event.

As shown in this study, the ORF's encoding SlaA and SlaB were successfully cloned into expression strains of *E. coli*. This was initially demonstrated with the successful amplification of *slaA* and *slaB* and successful release of each respective gene fragment following restriction enzyme digestion with appropriate restriction endonucleases. Short sequencing of the resulting construct also indicated that the sequence cloned represented the full mature protein and was shown to be intact.

As demonstrated, expression was not achieved using standard expression strains of *E. coli*. It can be speculated that a potential reason for this may be due to the inherent potential of secreted phospholipase enzyme to damage the bacterial host cell envelope thus severely affecting the survivability of *E. coli* expressing these constructs.

The work of Sitkiewicz and colleagues (Sitkiewicz *et al.* 2007) indicated difficulty in expressing constructs of wild type SpSlaA. The author attributed this to the cytotoxic effect of SpSlaA but suggested such effects required intracellular localization. It was also noted that expression of SpSlaA in yeast resulted in a rapid decrease in cell viability.

The analysis of the purified proteins obtained in this study also indicated that a substantial portion of purified SlaA protein aggregated, as indicated by Western –blot analysis using anti-His-Tag antisera. A disparity was also demonstrated between the detection of SlaA and SlaB; this however is likely due to differential amounts of protein added to each sample well.

The obtained quantities of purified proteins were not as favourable as expected, especially in the case of purified SlaB protein and as such did not permit protein crystallisation studies. It is unclear why the quantity of purified SlaB protein obtained was substantially lower than SlaA. SlaB also showed reduced stability compared to SlaA (Figure. 4.15). It should however be noted though that SlaA demonstrated a far lower protein instability index (23.83) than SlaB (36.45) when estimating protein stability for both mature proteins with the inclusion of a His₆-tag. A instability index of >40 is defined as ‘unstable’ by the method of Guruprasad and colleagues (Guruprasad *et al.* 1990).

Regarding the demonstrated activity of SlaA, a K_m of 14.40 ± 7.866 mM and a specific activity of $5.06 \times 10^{-2} \pm 3.01 \times 10^{-3}$ $\mu\text{mol}/\text{min}/\text{mg}$ were recorded. An obvious choice for comparison is the closely related, SpSlaA, characterised by Nagiec and colleagues (Nagiec *et al.* 2004). However despite the investigation of SpSlaA activity against a variety of substrates, no standard catalytic values were published. To my knowledge at time of writing no other sPLA₂ enzymes from a bacterial source have been cloned into an expression host system and tested for catalytic activity.

The substrate utilised in this study (DHT-PC) was also not the preferred substrate choice utilised in the non-colorimetric sPLA₂ enzymatic assays which were the reference standard until recently.

However comparisons can be made with the sPLA₂ standard used in this study obtained from *Apis mellifera*. Comparisons have shown that SlaA processed a specific activity 87.4% lower than the recorded specific activity of the standard sPLA₂. Considering that the standard sPLA₂ used has been shown to reversibly induce rabbit platelet aggregation at concentrations of 1-10 µg/ml (Ouyang and Huang 1984) it is highly likely that SlaA possesses a similar damaging effect *in vivo* at similarly relatively low concentrations. This potential *in vivo* exact level of expression of either *slaA* or *slaB* in *S. equi* and *slaB* in *S. zooepidemicus* however currently remains unknown. This could be revealed *in vivo* by the utilization of Reverse transcription PCR techniques. Moreover, the Western blot analysis of SlaA and SlaB against equine convalescent sera (Figure 4.17) is important as this demonstrates that SlaA is clearly presented to the equine host defences early in infections (due to a strong detection in strangles positive serum and no detection in strangles negative sera). Thus phospholipase activity of secreted SlaA may be a significant factor in damaging host cells during pathogenesis *in vivo*.

As demonstrated, SlaA was shown to have an optimum rate of activity with DHT-PC at 45 °C with a reduction of activity demonstrated at 50 °C. The significant standard deviation of the means at 50 °C perhaps suggest enzyme instability or degradation at a higher than optimum temperature. The standard body temperature of the healthy horse is 38°C, however as noted in section 1.3, the pathogenesis of an *S. equi* infection is strongly associated with an onset of fever (pyrexia). During equine fever the temperature of the horse can become

elevated to >40°C. The demonstration that SlaA remains potent enzymically at temperatures >40 °C could be considered significant with respect to the enzyme activity in situ.

SlaA displayed an enzymatic requirement for Ca^{2+} which is consistent with literature observations of other sPLA₂ enzymes as notated in BRENDA, the comprehensive enzyme information system (http://www.brenda-enzymes.org/php/result_flat.php4?ecno=3.1.1.4). The specificity of interaction with Ca^{2+} was also as expected, given the identification of a calcium binding loop (Figure 4.1), the ability of Mg^{2+} to act as a replacement of Ca^{2+} however was not. Some evidence of observations similar to the results demonstrated in this study regarding this have however been noted in the literature, most notably Mg^{2+} has been shown to replace Ca^{2+} with decreased activity in sPLA₂ of *Apis mellifera* (Cottrell 1981).

As demonstrated, no activity was shown for SlaB against either DHT-PC or AT-PC. A likely reason for this relates to the obtained quality and quantity of purified protein available for enzymatic assay. In the inaccurate Sigma/Price colorimetric assays reactivity was visually observed (data not shown). Regarding the Caymen assay, a small detectable optical absorbance measurement was also observed. However the level of absorbance changes between a standard reaction (standard 1.66 mM substrate concentration) of SlaB against DHT-PC and a blank standard was negligible and lacked significance. It could therefore be concluded that the lack of enzymatic activity displayed by SlaB was likely to have been a result of the low amount of enzyme available to assay rather than a significant biological characteristic.

It was highly unexpected that SlaA demonstrated no activity towards AT-PC since SpSlaA has previously been shown to release AA from phosphatidylcholine (Nagiec *et al.* 2004). Currently the amino acid residues in SlaA that differ from SpSlaA are also not known to play a role in enzyme catalysis. However this has proved difficult to confirm due to the

absence of a crystal structure of SlaA or SpSlaA and difficulty in structural predictions based on molecular models due to low sequence identities between both enzymes and structures deposited in the Protein Data Bank (Nagiec *et al.* 2004).

Several possible potential reasons can however be established for the observed lack of activity of SlaA against AT-PC. Firstly as demonstrated, activity was observed for a control sPLA₂ against AT-PC with a specific activity significantly 87.4% reduced from a control reaction against DHT-PC. Since SlaA was demonstrated to have a significantly lower specific activity against DHT-PC compared to that of the control sPLA₂ it is possible that the level of activity of SlaA against AT-PC may not have been at a sufficient level to be detected by the assay. A further explanation could relate to the biochemical nature of the assay as the colourimetric nature of the assay presumes that substrate cleavage is specific for the *sn*-2 position when using DHT-PC. Thus a cleavage at the *sn*-1 position would not be detected by this biochemical assay, if indeed SlaA fitted more accurately with sPLA₁ rather than sPLA₂ nomenclature.

However this is highly unlikely based on the observations of Nagiec and colleagues with SpSlaA which was shown to cleave substrates at the *sn*-2 position. Furthermore it has been established that SpSlaA is not a PLB enzyme, as incubation of SpSlaA with dipalmitoyl-[1, 2-palmitoyl-1-¹⁴C] PC produced detectable palmitic acid and LPC (Nagiec *et al.* 2004) so hence SlaA clearly does not have the ability to cleave both *sn*-1 and *sn*-2 positions.

Regarding future work, in order to provide strong evidence regarding SlaA and SlaB as bona fide virulence factors it is necessary to produce a isogenic mutant strains of *S. equi*, one with an inactivated *slaA* gene and another with an inactivated *slaB* gene, and also a *slaA/slaB* double mutant (with relevant complementation strains). This was attempted in this

study utilising the same methodology as described in section 3.8. Unfortunately at time of writing this has not yet been successful. It would however be expected that Δ SlaA would also possess a similar reduced ability to adhere to equine epithelial cells as noted in SpSlaA (Sitkiewicz *et al.* 2007), unless there is redundancy between SlaA and SlaB (which could be determined by comparison with a double mutant). Comparing Δ SlaB and Δ SlaASlaB mutants would provide the ability to assess the biological comparisons between *S. equi* and *S. zooepidemicus* and provide potential clues to the molecular basis of differing pathogenicity amongst the two organisms.

As previously demonstrated in figure 4.2, SlaA closely resembles SpSlaA with two amino acid substitutions (At amino acid positions 34 and 108). However as Nagiec and colleagues demonstrated in *SpSlaA* gene mutagenesis experiments, even single substitutions were shown to result in either minor or significant reductions in enzyme activity or in some cases an increase in activity (Nagiec *et al.* 2004). As expected, substitutions in critical regions (Such as the catalytic site and Ca^{2+} binding loop) resulted in a significant reduction of activity. However of particular interest is the substitution of the tyrosine at amino acid position 101 which resulted in a 50% loss in enzyme activity.

Of the amino acid substitutions conducted by Nagiec and colleagues neither of the residues corresponding to those unique to SlaA was investigated (Figure 4.28). The tyrosine at position 101 is however situated in a portion of the protein predicted to possess a coil secondary structure (data not shown), in a region shared by a SlaA variant amino acid (an isoleucine at position 108). It is therefore appealing to suspect that if a substitution in an amino acid in SpSlaA not associated with any critical region of the protein primary structure can result in a drastically different rate of enzyme activity then a variant amino acid in SlaA in a similar position may also have a distinguishing effect on the properties of the enzyme. In

the absence of a crystal structure of SlaA the production of *SlaA* gene mutants possessing substitutions at amino acid position 34 and 108 by site-directed mutagenesis would allow this investigation.

EGINDKIENGTESDISFQNGELLKNYLI LEGERVYFDYDRATQDKVS
DDVLEMGMLVEAISKDYSEKTFTPDKYFKASWP_IHGNYCGPGHNGNN
FTLPVVDVLDQGCQNHDSCKWGAGIGANCECNRQLVNYIKVYRRWM
PANVLGVADAIRVYFETVGSIGC

Figure 4.28 Amino acid sequence of SlaA highlighting the effect of amino acid substitution as demonstrated by Nagiec and colleagues (Nagiec *et al.* 2004).

Amino acids unique to SlaA and absent in SpSlaA are underlined. Amino acids in red have a significant reduction in enzyme activity when substituted (below 25% of total activity). Amino acids in blue have moderate to little reduction in enzyme activity when substituted (between 25-100% of total activity). Amino acids in green have a small increase in enzyme activity when substituted (between 0-50% increases of total activity). Amino acid in yellow have a significant increase in enzyme activity when substituted (over 50% increase of total activity).

5. Investigation into the Acid phosphatase, LppC of *S. equi* and *S. zooepidemicus*

5.1. Introduction

As previously described in Section 1.9, acid phosphatases are representative of a group of enzymes that may participate in the virulence of pathogenic bacterial species with phosphatase activity indicated to be highly variable between even closely related species. Western blotting using rabbit polyclonal antibody to the Class C NSAP, LppC has previously indicated the presence of acid phosphatases in *S. equi* strains 1026, 1742, 4047, 9682 and *S. zooepidemicus* strains 7023, 3682, 461 and K3 using whole cell extracts (Hamilton *et al.* 2000). The non-biologically significant substrate *p*NPP was also employed to demonstrate acid phosphatase activity in *S. equi* 9682 using whole cell extracts. One acid phosphatase activity was detected which possessed an activity optimum at pH 5.0 and another optimum at 6.0-6.5, with only the former also exhibiting resistance to EDTA inhibition (similar to LppC of *S. equisimilis*).

Zymographic analysis of *S. equi* strains 9682 and 4047 also detected a low level of acid phosphatase activity; the low level of zymographic activity has been attributed to the use of whole cell extracts during analysis.

PCR was originally used to identify an internal sequence (340 bp) of the *S. equi* and *S. zooepidemicus* *lppC* genes (Hamilton *et al* 2000). With the availability of the *S. equi* and *S. zooepidemicus* genomes, it has now been possible to identify *seq0346*, the gene responsible for encoding LppC in *S. equi* (herein referred to as *SeLppC*). Furthermore with the availability of the *S. zooepidemicus* genome it has also been possible to bioinformatically

identify an orthologue of *SeLppC* in *S. zooepidemicus* H70, notated as *szo16870* (herein referred to as *SzLppC*). However, the *S. zooepidemicus* H70 genome sequence indicates a possible frame shift in the *SzLppC*-5' coding region that would introduce a stop codon in the *SzLppC* lipoprotein signal peptide sequence. Thus *szo16870* was annotated as pseudogene in *S. zooepidemicus* H70 (Holden *et al* 2009) and so LppC may represent an additional difference between *S. equi* and *S. zooepidemicus*

Currently, the catalytic rate of activity of *SeLppC* and *SzLppC* remains to be determined. The substrate specificity and other physicochemical parameters of *SeLppC* and *SzLppC* also remain to be determined, with other with NSAP's described as retaining activity across a wide variety of substrates as described in section 1.9.1.

Aims of the present study

The aims of the present study were:

- a) To assess the presence of putative class C NSAP in the genome sequenced strains of *S. equi* and *S. zooepidemicus* using Western blotting of whole cell extracts.
- b) To short sequence *SzLppC* to investigate the possible frame shift in the *SzLppC* coding region.
- c) To clone, *SeLppC* and *SzLppC* identified in *S. equi* 4047 and *S. zooepidemicus* H70 respectively for expression and purification from *E. coli*.
- d) To detect recombinant forms of *SeLppC* and *SzLppC* with zymography
- e) To assess the enzymatic properties of *SeLppC* and *SzLppC* and provide a comparison to previously characterised similar class C NSAP enzymes.

5.2. Bioinformatic analysis of SeLppC and SzLppC

The nucleotide and amino acid sequences of the putative SeLppC (SEQ0346) and SzLppC (SZO16870) were obtained from Uniprot in a FASTA format. Multiple alignment of the two full nucleotide sequences indicated similarity of 99% (698/704), highlighting the possible frame shift in the *SzLppC*-5' coding region between nucleotides 15 and 16 of *SzLppC* (Figure 5.1). Full translation of the nucleotide sequence of *SzLppC* does indicate that *SzLppC* is a putative pseudogene with a stop codon introduced at nucleotide position 3 (Figure 5.2). It should be noted that *SzLppC* is not listed as a pseudogene in the *S. zooepidemicus* MGAS 10565 genome.

The sequence encoding SeLppC was analysed for identification of a lipoprotein signal peptide cleavage site by using Pred-Lipo Hidden Markov Models outputs, with a cleavage site predicted between 21 and 22 with a probability of 0.994. A tBLASTn search of the translated *S. equi* genome database for homologues to SeLppC indicated no significant homologues other than itself. Multiple alignment of the mature amino acid sequences of SeLppC and SzLppC reveal that both proteins are nearly identical with the exception of an amino acid substitution at position 2 of the mature protein (Figure 5.3). Database searches for orthologues to the amino acid sequence of the mature SeLppC protein using BLAST search revealed a large number of proteins with significant similarities (Table 5.1).

SeLppC	ATGACATCTAAAAAAATTCTTAACCTTGTGTCTTTAGGGTTATCGTTGGTTTTAATCTCT	60
SzLppC	ATGACATCTAAAAAA- <u>TT</u> CTTAACCTTGTGTCTTTAGGGTTATCGTTGGTTTTAATCTCT	59

SeLppC	GGCTGT <u>AC</u> GACAAATGGAGAGAAAAAAGCAGCTCCTAGTCAAGACAAGGATAAGCAAGAA	120
SzLppC	GGCTGT <u>G</u> CGACAAATGGAGAGAAAAAAGCAGCTCCTAGTCAAGACAAGGATAAGCAAGAA	119

SeLppC	AAGGTTGTCCGTCTAACCAATGATCAGCTAAGGGCTAGGGAAAATACCATGGCCACACTC	180
SzLppC	AAGGTTGTCCGTCTAACCAATGATCAGCTAAGGGCTAGGGAAAATACT <u>AT</u> GGCCACACTC	179

SeLppC	TGGTATCAGCAGTCTGAAGAGGCTAAGGCTCTCTACTTACAGGGGTACCAAGTGGCTAAG	240
SzLppC	TGGTATCAGCAGTCTGAAGAGGCTAAGGCTCTCTAC <u>CT</u> ACAGGGGTACCAAGTGGCTAAG	239

SeLppC	CAGCGCTTAGATACATTATTGAGTCAAGCAACTGATAAACCATACTCTATTGTTTTGGAC	300
SzLppC	CAGCGCTTAGATACATTATTGAGTCAAGCAACTGATAAACCATACTCTATTGTTTTGGAC	299

SeLppC	ATTGATGAGACAGTGCTCGATAAATAGTCCTTACCAGGCTAAAAATATTAAAGAGGGCACA	360
SzLppC	ATTGATGAGACAGTGCTCGATAAATAGTCCTTACCA <u>AG</u> CTAAAAATATTAAAGAGGGCACA	359

SeLppC	GGCTTTACGCCGGATTCTTGGGATAAATGGGTCCAAAAGAAATCAGCCAAGGCAGTAGCA	420
SzLppC	GGCTTTACGCCGGATTCTTGGGATAAATGGGTCCAAAAGAAATCAGCCAAGGCAGTAGCA	419

SeLppC	GGTGCTAAGGACTTCTTGCAAGTATGCAAATGACAAGGGCGTTCAGATTTACTATGTTTCT	480
SzLppC	GGTGCTAAGGACTTCTTGCAAGTATGCAAATGACAAGGGCGTTCAGATTTACTATGTTTCT	479

SeLppC	GACAGAACGACTAAGCAAGTAGAGCCGACTATGGAAAATCTTGAAAAAGAAGGTATTCCA	540
SzLppC	GACAGAACGACTAAGCAAGTAGAGCCGACTATGGAAAATCTTGAAAAAGAAGGTATTCCA	539

SeLppC	GTGCAAGGCAAGGATCATTTCTTGTTCTTAGAGGAGGGTGTA ^{AA} ATCAAAGAGGGCCGT	600
SzLppC	GTGCAAGGCAAGGATCATTTCTTGTTCTTAGAGGAGGGTGTA ^{AA} ATCAAAGAGGGCCGT	599

SeLppC	CGTCAAAGGTGCAAGAAACAACA ^{AA} ATCTAGTGTTGTTATTTGGTGACA ^{AT} TTTGCTTGAT	660
SzLppC	CGTCA ^{AA} AGGTGCAAGAAACAACA ^{AA} ATCTAGTGTTGTTATTTGGTGAT ^A ATTTGCTTGAT	659

SeLppC	TTTGCAGAGTTTTCTAAGACCTCTCATGAAGATAGAAGAAAGCTT	705
SzLppC	TTTGCAGAGTTTTCTAAGACCTCTCATGAAGATAGAAGAAAGCTT	704

Figure 5.1: Pairwise alignment of *SeLppC* and *SzLppc*.

Alignment was performed using ClustalW2. * indicates identical nucleotides. The variant nucleotide positions are underlined. The sequences start at the signal peptide region. Stop position introduced by the frameshift (missing A from the poly A tract) is underlined.

Met T S K K F T L C L Stop

K I L N L V S L G L S L V L I S G C A T N G E K K A A
P S Q D K D K Q E K V V R L T N D Q L R A R E N T **Met**
A T L W Y Q Q S E E A K A L Y L Q G Y Q V A K Q R L D
T L L S Q A T D K P Y S I V L D I D E T V L D N S P Y
Q A K N I K E G T G F T P D S W D K W V Q K K S A K A
V A G A K D F L Q Y A N D K G V Q I Y Y V S D R T T K
Q V E P T **Met** E N L E K E G I P V Q G K D H F L F L E
E G V K S K E G R R Q K V Q E T T N L V L L F G D N L
L D F A E F S K T S H E D R R K L

Figure 5.2: Translated amino acid sequence of SzLppC, either side of the stop codon introduced by the putative frameshift

Translation was undertaken using Expasy translate tool. Stop position introduced by the frameshift (missing A from the poly A tract) is underlined.

SeLppC	CTTNGEKKAAPSQDKDKQEKVVRLTNDQLRARENTMATLWYQQSEEAKALYLQGYQVAKQ	60
SzLppC	CATNGEKKAAPSQDKDKQEKVVRLTNDQLRARENTMATLWYQQSEEAKALYLQGYQVAKQ	60
	* : *****	
SeLppC	RLDTLLSQATDKPYSIVLDIDETVLDNSPYQAKNIKEGTGFTPDSDKWVQKKSAAVAG	120
SzLppC	RLDTLLSQATDKPYSIVLDIDETVLDNSPYQAKNIKEGTGFTPDSDKWVQKKSAAVAG	120

SeLppC	AKDFLQYANDKGVQIYYVSDRTTKQVEPTMENLEKEGIPVQGKDHFLFLEEGVKSKEGRR	180
SzLppC	AKDFLQYANDKGVQIYYVSDRTTKQVEPTMENLEKEGIPVQGKDHFLFLEEGVKSKEGRR	180

SeLppC	QKVQETTNLVLLFGDNLLDFAEFSKTSBEDRRKLLDQLHAEFGSKFIIFPNPMYGSWESS	240
SzLppC	QKVQETTNLVLLFGDNLLDFAEFSKTSBEDRRKLLDQLHAEFGSKFIIFPNPMYGSWESS	240

SeLppC	VYNGQKLDSEGQIKARDKALEAY	263
SzLppC	VYNGQKLDSEGQIKARDKALEAY	263

Figure 5.3: Pairwise alignment mature of SeLppC and SzLppc.

Alignment was performed using ClustalW2. * indicates identical amino acids, : indicates similar amino acids. The variant amino acid position is underlined. The sequences start from the lipidated cysteine of the mature lipoprotein.

Table 5.1: Significant orthologues of mature SeLppC. Blast search performed by tBLASTn

<u>Accession</u>	<u>Annotation</u>	<u>Organism</u>	<u>Identities</u>	<u>Positives</u>	<u>Gaps</u>	<u>E number</u>
gb ABV10563.1	5'-nucelotidase, lipoprotein, e(P4) family	<i>Streptococcus gordonii</i> str. Challis substr. CH1	174/242 (72%)	210/242 (87%)	0/242 (0%)	9.00E-133
gb ALL98443.1	Acid phosphatase	<i>Streptococcus pyogenes</i> MGAS8232	169/239 (71%)	203/239 (85%)	0/239 (0%)	1.00E-126
gb EFR44871.1	5'-nucelotidase, lipoprotein, e(P4) family	<i>Streptococcus pseudoporcinus</i> SPIN 20026	185/268 (69%)	217/268 (81%)	11/268 (4%)	2.00E-133
YP_002997631.1	5'-nucelotidase, lipoprotein, e(P4) family	<i>Streptococcus dysgalactiae</i> subsp. <i>equisimilis</i> GGs_124	182/263 (69%)	213/263 (81%)	1/263 (0%)	3.00E-132
gb EDT27341.1	5'-nucleotidase, lipoprotein e(P4) family	<i>Clostridium perfringens</i>	99/211 (43%)	137/211 (65%)	3/211 (1%)	3.00E-58
gb EEW75817.1	Lipoprotein e(P4) family 5'-nucleotidase	<i>Haemophilus influenzae</i> RdAW	79/229 (34%)	124/229 (54%)	8/229 (3%)	4.00E-35

5.3. PCR across *SzLppC* Signal Peptide region

In order to short sequence the region of interest the designed primer pair, H70APSEQFOR and H70APSEQREV (Table 2.4) was used to amplify one PCR fragment from the genomic DNA of *S. zooepidemicus*. The primer set was designed to amplify a short section of DNA, spanning the nucleotide region suspected of representing the possible frame shift in the SzLppC signal peptide coding sequence. Amplification of the primer set yielded a product consistent with the expected length of 163bp (Figure 5.4)

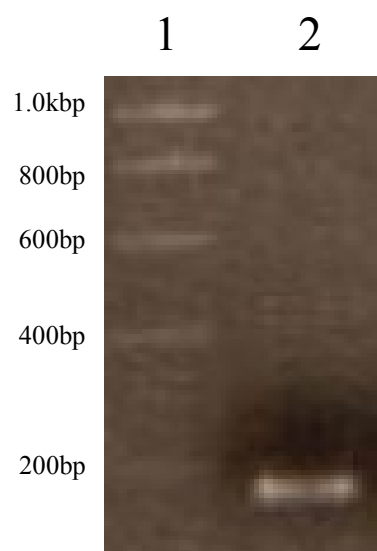


Figure 5.4: PCR amplification of *SzLppC* signal peptide region.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standard; Lane 2, PCR product from amplification of *SzLppC* signal peptide coding region.

5.4. DNA sequencing of *SzLppC*

The product amplified from the designed primer pair, H70APSEQFOR and H70APSEQREV (Figure 5.4) was subjected to Sanger Sequencing. In total 4 replicate samples were sequenced and the obtained sequences were aligned to confirm sequence integrity. The background trace was also used to confirm the obtained sequence (see appendix IV). The obtained partial sequence for *SzLppC* was then compared to that of the annotated sequence obtained from the *S. zooepidemicus* H70 genome project (Figure 5.5).

ClustalW alignment of the obtained sequence gave 99% positive identity (103/104 conserved nucleotides) between the sequenced region of *SzLppC* and corresponding sequence in the *S. zooepidemicus* H70 genome. The gap indicated in the *S. zooepidemicus* H70 genome database between the thymine at position 1887168 and the adenine at position 1887169 was shown to be incorrect with the obtained sequence of *szo16870* shown to have a correctly positioned thymine at position 1887169 of the genome. Note that *SzLppC* is encoded on the reverse strand and so the sequence presented in Figure 5.1 is the reverse complement of that in Figure 5.5.

These data indicate that the deposited sequence for *S. zooepidemicus* H70 contains a sequencing error and that the true *szo16870* sequence is not a pseudogene. This is further supported by the observation that the corrected *S. zooepidemicus* H70 sequence is now identical to the corresponding region of the *S. zooepidemicus* MGCS10565 genome sequence.

```

Query:      2 AAACCAACGATAACCCTAAAGACACAAGGTTAAGAATTTTTTTTAGATGTCATAAAATCCT 61
              |||
Sbjct: 1887127 AAACCAACGATAACCCTAAAGACACAAGGTTAAGAATTTTTT-AGATGTCATAAAATCCT 1887185

Query:      62 CCTAATATTCCATACATTATTAAGTACGGTTATTTATCCACCTC 105
              |||
Sbjct: 1887186 CCTAATATTCCATACATTATTAAGTACGGTTATTTATCCACCTC 1887229

```

Figure 5.5: Multiple alignment of sequenced portion of *SzLppC* (this study, 4 replicates) and the deposited nucleotide sequence of *SzLppC*, obtained from the *S. zooepidemicus* H70 genome.

Multiple alignment was performed using ClustalW2. ‘Query’ represents obtained *SzLppC* sequence and ‘sbjct’ represents nucleotide sequence of *SzLppC*, obtained deposited *S. zooepidemicus* H70 genome sequence.

5.5. Identification of LppC enzymes in *S. equi* and *S. zooepidemicus*

The presence of acid phosphatase lipoprotein in a range of whole cell protein extracts of *S. equi* and *S. zooepidemicus* strains (Figure 5.6) was screened by Western immunoblotting, using a polyclonal antiserum raised against LppC of *S. equisimilis* (Malke 1998). The antiserum was shown to cross-react strongly at a molecular weight consistent with the expected size of *SeLppC* (32 kDa) with three strains of *S. equi*, i.e 4047 and 1026 (as shown by Hamilton et al 2000) and 9672. Cross reaction was also evident with four *S. zooepidemicus* strains; (7626, H70, 3635 and K3) but detection level was significantly reduced. Of these strains, only K3 has previously been investigated for LppC (Hamilton et al. 2000). Notably antibody-cross reactivity was not detected in *S. zooepidemicus* 3682 and was detected strongly in strain *S. zooepidemicus* H70 but at a reduced molecular weight of under 30 kDa (Figure 5.7). Three replicate growths of the *S. zooepidemicus* strains were undertaken and this observation was confirmed throughout the replicates (data not shown).

This data is consistent with the resequencing data indicating that *S. zooepidemicus* H70 specifically (and *S. zooepidemicus* strains generally) are capable of expressing LppC i.e. that the annotation of *szo16870* as a pseudogene in the *S. zooepidemicus* H70 genome is a sequencing artefact.

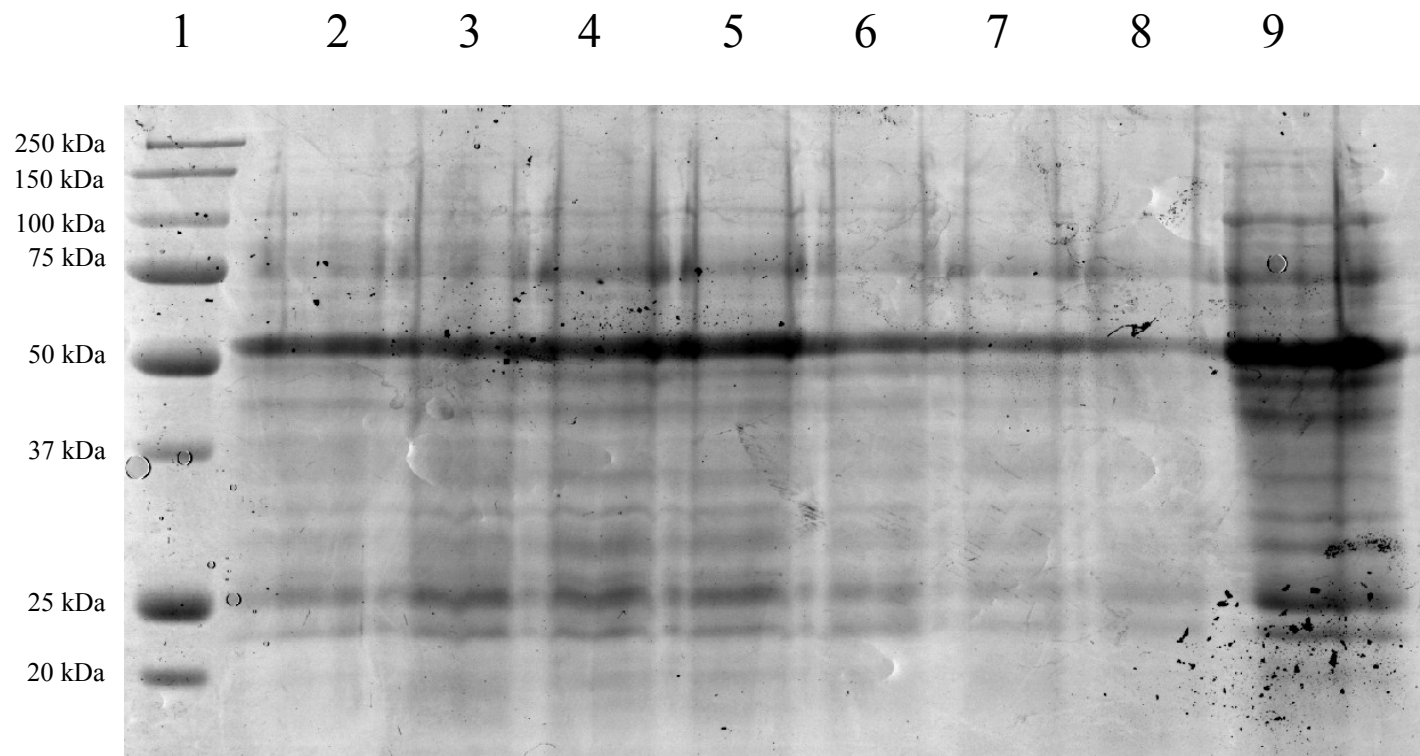


Figure 5.6: 15% SDS PAGE of range of *S. equi* and *S. zooepidemicus* strains.

Lane were loaded as follows: Lane 1, molecular weight markers (250, 150, 100, 75, 50, 37, 25, 20 kDa); lane 2, *S. equi* 9672; lane 3, *S. equi* 4047; lane 4, *S. equi* 1026; lane 5, - *S. zooepidemicus* 7622; lane 6, - *S. zooepidemicus* H70; lane 7, *S. zooepidemicus* 3635; lane 8, *S. zooepidemicus* K3; Lane 9, *S. zooepidemicus* 3682.

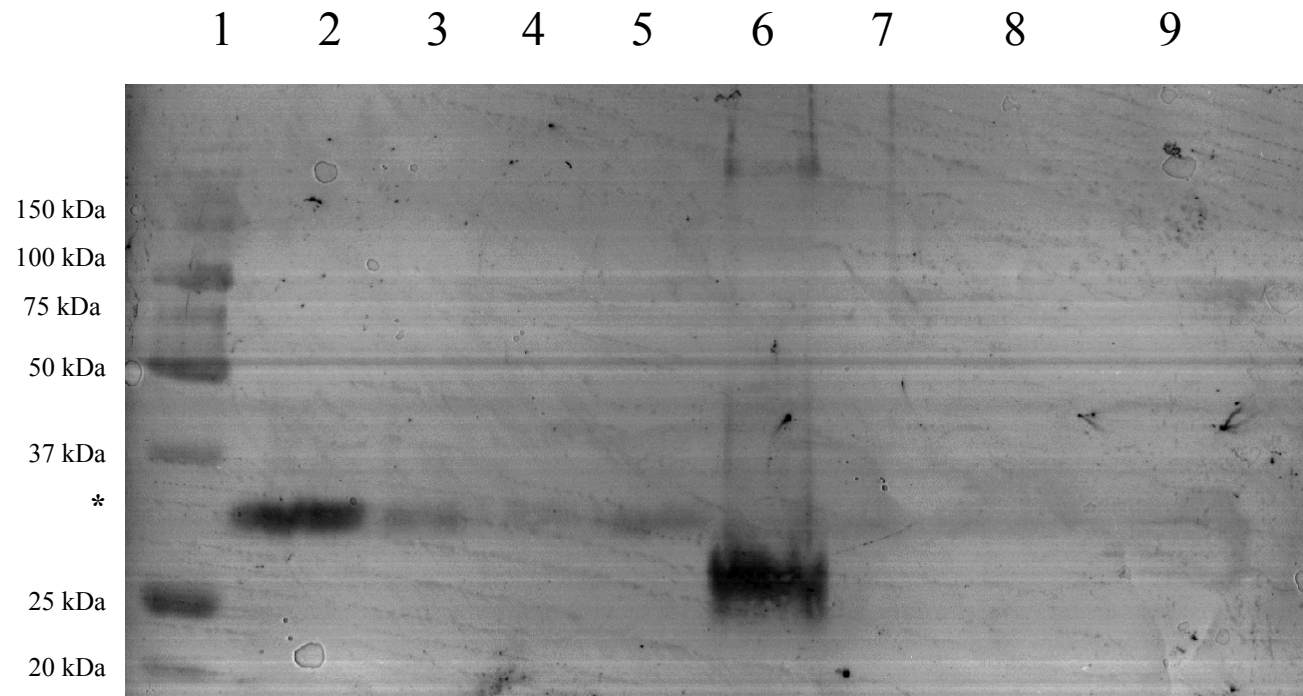


Figure 5.7: LppC Western blot analysis of a range of *S. equi* and *S. zooepidemicus* strains.

Lanes were loaded as follows: Lane 1, molecular weight markers (250, 150, 100, 75, 50, 37, 25, 20 kDa); lane 2, *S. equi* 9672; Lane 3, *S. equi* 4047; lane 4, *S. equi* 1026; Lane 5, *S. zooepidemicus* 7622; lane 6, *S. zooepidemicus* H70; lane 7, *S. zooepidemicus* 3635; lane 8, *S. zooepidemicus* K3; lane 9, *S. zooepidemicus* 3682.

5.6. PCR of mature *SeLppC* and *SzLppC*

Using the designed primers pair, SEQLPPCf- SEQLPPCr (Table 2.4), two PCR fragments were amplified from the genomic DNA of *S. equi* and *S. zooepidemicus* (Figure 3.4 and Figure 3.5 respectively). Both primer sets amplified only a portion of each representative genome encompassing the necessary gene product that encoded the mature SeLppC/SzLppC protein and a small non-coding region downstream of the target genes. Amplification of the primer set yielded two products consistent with their predicted length of 817 bp (Figure 5.8). Notably these primers also amplified additional larger PCR products which were not investigated further as BLAST searches strongly suggest that *seq0346* and *szo16870* are the only putative LppC homologues encoded in the *S. equi* 4047 genome and *S. zooepidemicus* H70 genome respectively.

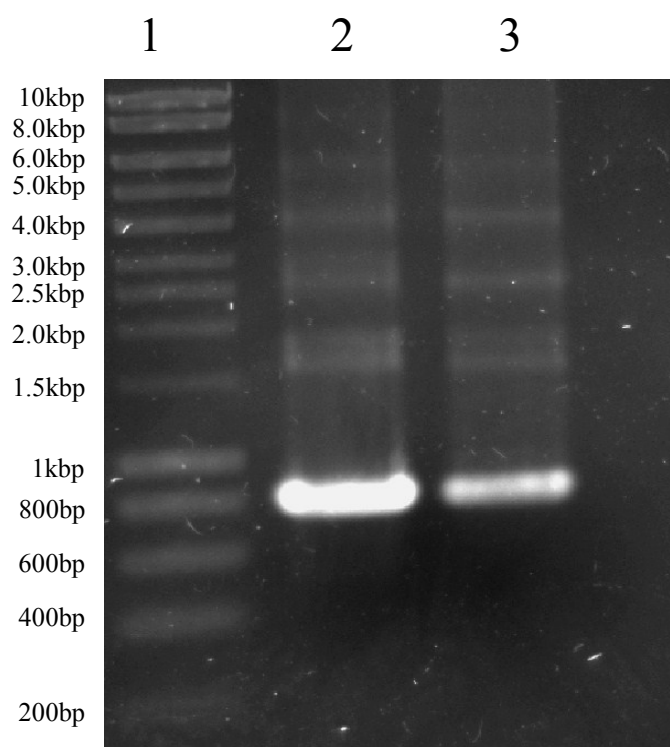


Figure 5.8: PCR amplification of *seq0346* and *szo16870*.

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standard; Lane 2, PCR product from amplification of *seq0346* from *S. equi* 4047 genomic DNA; Lane 3, PCR product from amplification of *szo16870*.

5.7. Molecular cloning of *SeLppC* and *SzLppC*

The amplified DNA products obtained from the designed primer pair, SEQLPPCf-SEQLPPCr (Figure 5.8) were initially cloned into pCR[®]-Blunt vector, yielding plasmids pbSeLppC and pbSzLppC. These plasmids were subsequently digested with appropriate restriction endonucleases to verify product integration. Cloned target genes and pCR[®]-blunt plasmid DNA was then digested with appropriate endonucleases, purified and cloned into similarly restriction digested and purified pET 28a vector. This generated plasmids pSeLppC and pSzLppC that encoded the proteins of interest with a His₆-tag at the N-terminus of the mature protein. Confirmation of the correct plasmid construction was undertaken by screening of plasmids extracted from candidate positive clones with appropriate restriction endonucleases. Following digestion, a successful clone was indicated by the visualisation of a product consisting of the target gene size and the remaining restriction endonuclease cut pET 28a vector, 5.4 kbp in size (Figure 5.9). Further confirmation was undertaken by the gene sequencing of the plasmid T₇ promoter region. The sequences obtained were checked for fidelity and aligned to their respective nucleotide sequences in the *S. equi* 4047 and *S. zooepidemicus* H70 genome (See appendix IV). Technical assistance from Mian Bilal Munir with this cloning work is acknowledged.

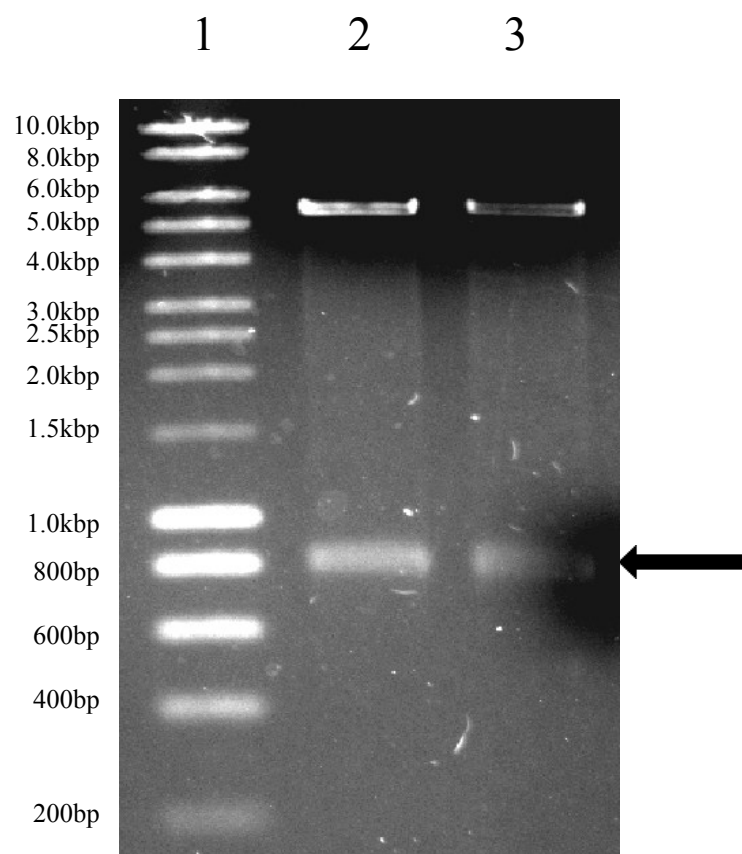


Figure 5.9: Restriction endonuclease digestion of pSeLppC and pSzLppC

Samples were run on a 1% agarose gel containing ethidium bromide (0.5 µg/ml). Lanes were loaded as follows: Lane 1, 10 kbp size standards; lane 2, pSeLppC post treatment with *EcoRI* and *XhoI* restriction endonucleases; lane 3, pSzLppC post treatment with *EcoRI* and *XhoI*. Released *SeLppC* and released *SzLppC* product (expected sizes of 817 bp) are indicated with an arrow.

5.8. Protein expression of SeLppC and SzLppC

Plasmids pSeLppC and pSzLppC encoding *SeLppC* and *SzLppC* respectively were chemically transformed into a range of *E. coli* expression strains (Table 2.1) and grown at 37 °C to mid-exponential phase. Initially, expression of the recombinant proteins was induced by the addition of 1 mM IPTG (final concentration) and the cultures were incubated at 20°C with two replicates also incubated at 30°C and 37°C (Figure 5.10).

All standard *E. coli* strains tested yielded poorly expressed target proteins. The presence of the *pLysE* plasmid is thought to potentially improve expression of soluble and active recombinants, by providing additional stability to target genes via the inhibition of T7 RNA polymerase (Studier 1991) . Chemical transformation of pSeLppC and pSzLppC into *E. coli* BL21 (DE3) *pLysE* cells was shown to significantly improve the expression of recombinant protein sufficient for further studies. It was also shown that variation in incubation temperature between 20°C and 37°C produced no notable variation (Figure 5.11).

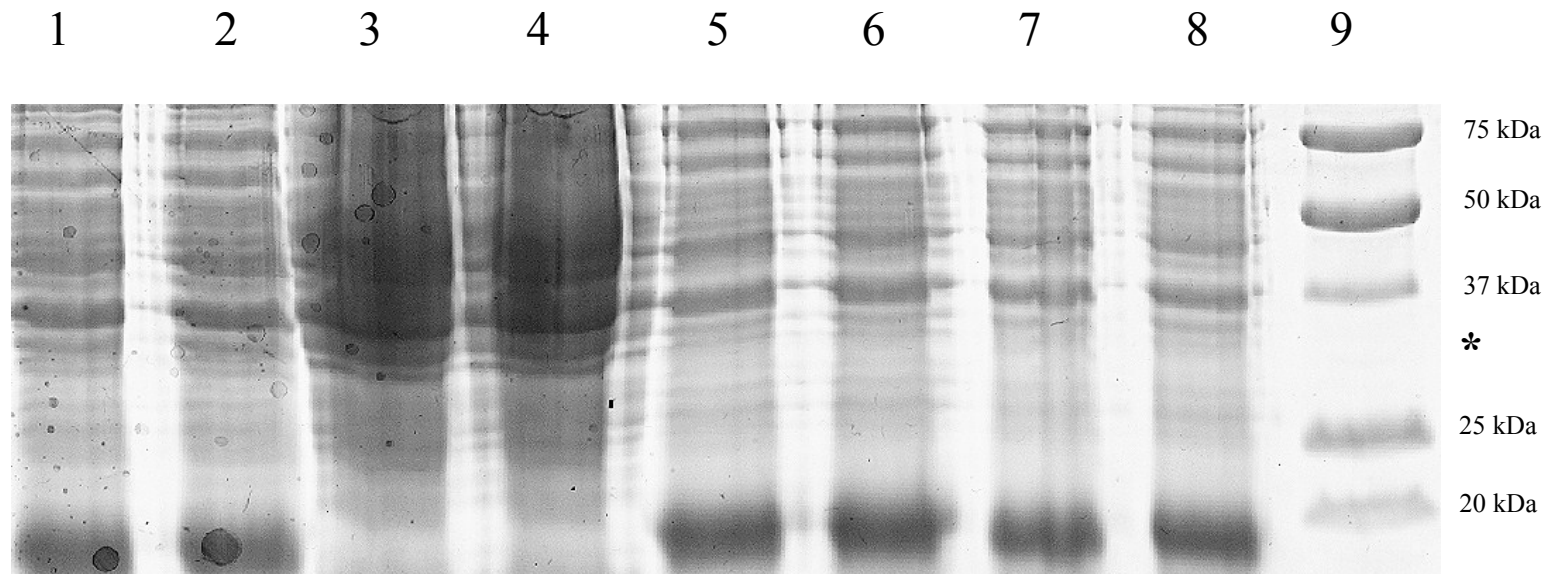


Figure 5.10: 15% SDS PAGE gel demonstrating SeLppC and SzLppC expression trials.

Protein expression was induced by 1 mM IPTG and growth at 20°C. The expected position of both proteins is marked with an asterisk. Lanes were loaded as follows: Lane 1, *E. coli* BL21 (DE3) pSeLppC; lane 2, *E. coli* BL21 (DE3) pSzLppC; lane 3, *E. coli* C41 (DE3) pSeLppC; lane 4, *E. coli* C41 (DE3) pSzLppC; lane 5, *E. coli* B834 (DE3) pSeLppC; lane 6, *E. coli* B834 (DE3) pSzLppC; lane 7, *E. coli* Tuner (DE3) pSeLppC; lane 8, *E. coli* Tuner (DE3) pSzLppC; lane 9, molecular weight markers (20, 25, 37, 50, 75 kDa).

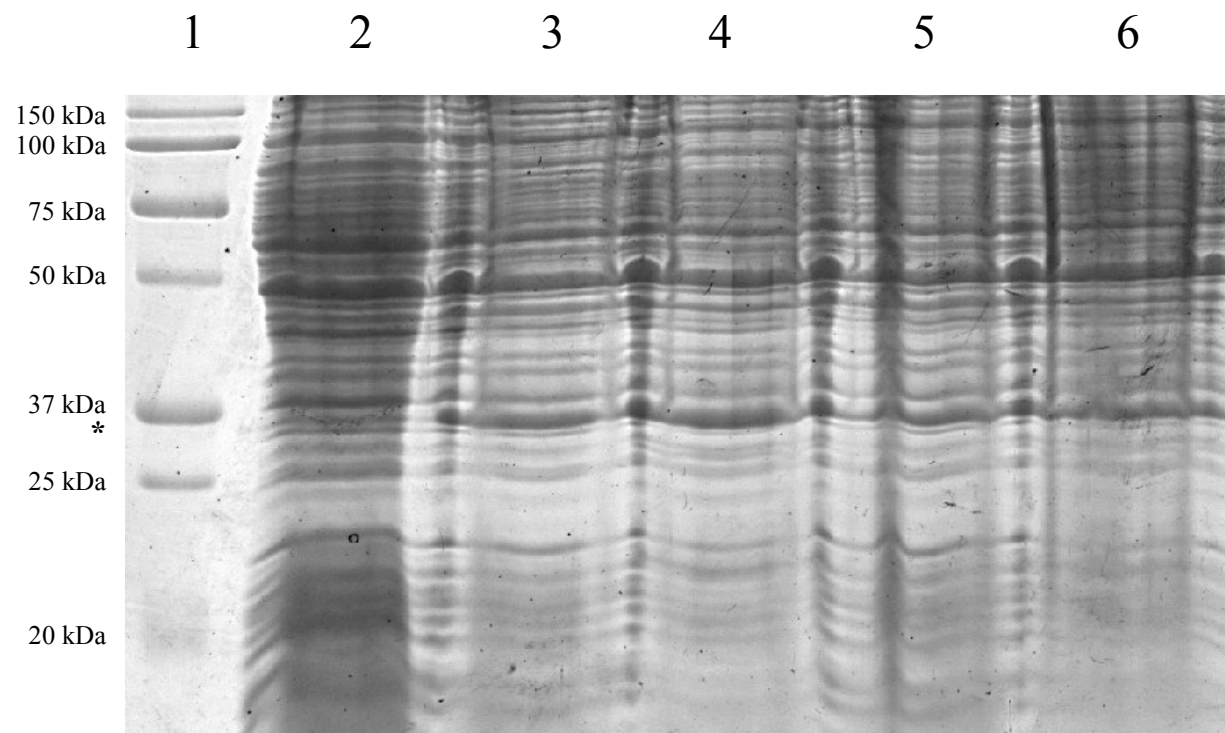


Figure 5.11: 15% SDS PAGE demonstrating SeLppC and SzLppC expression.

Protein expression was induced by 1 mM IPTG with *E. coli* (DE3) carrying the *pLysE* plasmid. The expected position of both SeLppC and SzLppC is marked with an asterisk. Lanes were loaded as follows: Lane 1, molecular weight markers (150, 100, 75, 50, 37, 25, 20 kDa); lane 2, unexpressed pSeLppC plasmid; lane 3, pSzLppc at 20°C; lane 4, pSzLppC at 37°C; lane 5, pSeLppC at 20°C; lane 6, pSeLppC at 37°C

5.9. Protein purification of SeLppC and SzLppC

Both His₆ –tagged proteins were expressed during *E. coli* growth in 4 litres of LB media and purified from cell free extracts by IMAC using a chelating Sepharose column and an automated gradient elution technique (Section 2.13.4). Confirmation of proteins present in different fraction was indicated by U.V absorption; fractions were checked for purity by SDS-PAGE and concentrated. Proteins were then further purified by FPLC, monitored by U.V absorption and checked for purity by SDS-PAGE, concentrated and quantified. Quantities of purified SeLppC and SzLppC obtained were ~0.625 mg/L.

A summary of all proteins produced for use in further studies is shown in Figure 5.12, which also shows the purity of a control commercial acid phosphatase from potato.

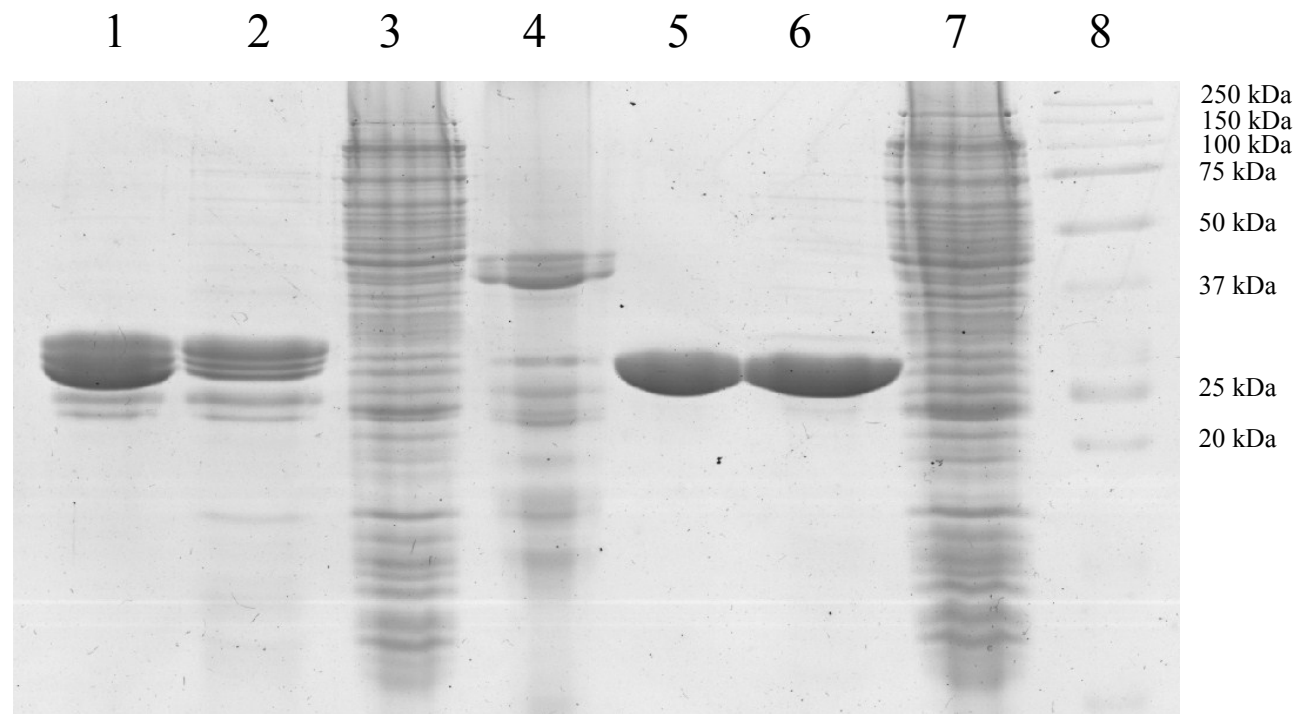


Figure 5.12: 15% SDS PAGE demonstrating proteins used in further work in this study.

Lanes were loaded as follows: lane 1, SeLppC (FPLC purified protein); lane 2, SeLppC (IMAC purified protein); lane 3, SeLppC (Cell free extract); lane 4, Acid phosphatase (Source: Potato); lane 5, SzLppC (FPLC purified protein); lane 6, SzLppC (IMAC purified protein); lane 7, SzLppC (Cell free extract); lane 8, molecular weight markers (250, 150, 100, 75, 50, 37, 25, 20 kDa).

5.10. Confirmation of His₆-tagged SeLppC and SzLppC

The presence of His₆-Tagged SeLppC and SzLppC in the protein samples obtained was screened by Western immunoblotting, using a polyclonal antiserum raised against the His₆-tag. The antiserum was shown to cross-react with proteins consistent with the size expected for SeLppC and SzLppC (32 kDa) in both the FPLC purified SeLppC and SzLppC products (Figure 5.13). A whole cell extract of *S. equi* was used as a negative control and showed no cross-reaction with the anti-His antibody. The molecular marker utilised was unstained and was only visible using reversible staining with Ponceau S reagent (marked in pencil prior to western blot development).

5.11. Detection of SeLppC and SzLppC using anti-LppC antiserum

The obtained purified protein products of SeLppC and SzLppC were screened by Western immunoblotting, using the polyclonal antiserum raised against LppC of *S. equisimilis* along with two control whole cell samples of each organism. The antiserum was shown to cross-react strongly with SeLppC purified protein product, additionally whole cell extracts of *S. equi* (strains 4047 and 9672) also cross-reacted similarly to that demonstrated in Figure 5.7. Only very weak cross-reactivity was shown with whole cell extracts of *S. zooepidemicus* (H70 and 3682), However the SzLppC purified protein product demonstrated a strong cross-reactivity similar to that demonstrated by SeLppC (Figure 5.14).

5.12. Western immunoblotting of purified SeLppC and SzLppC using equine convalescent serum

The presence of antibodies to SeLppC and SzLppC in equine convalescent serum was investigated by Western immunoblotting, using equine convalescent serum acquired from three ponies, pre and post infection with *S. equi*. Very faint antibody cross-reactivity with SeLppC and SzLppC was detected in all sera acquired post-infection, although only poor quality blot images could be obtained (Figure 5.15). No cross-reactivity was detected in pre-infection serum acquired from the ponies (not shown). Note that two different molecular markers were used in each separate panel. Both panels utilised molecular markers that were unstained and were only visible using reversible staining with Ponceau S reagent (marked in pencil prior to Western blot development).

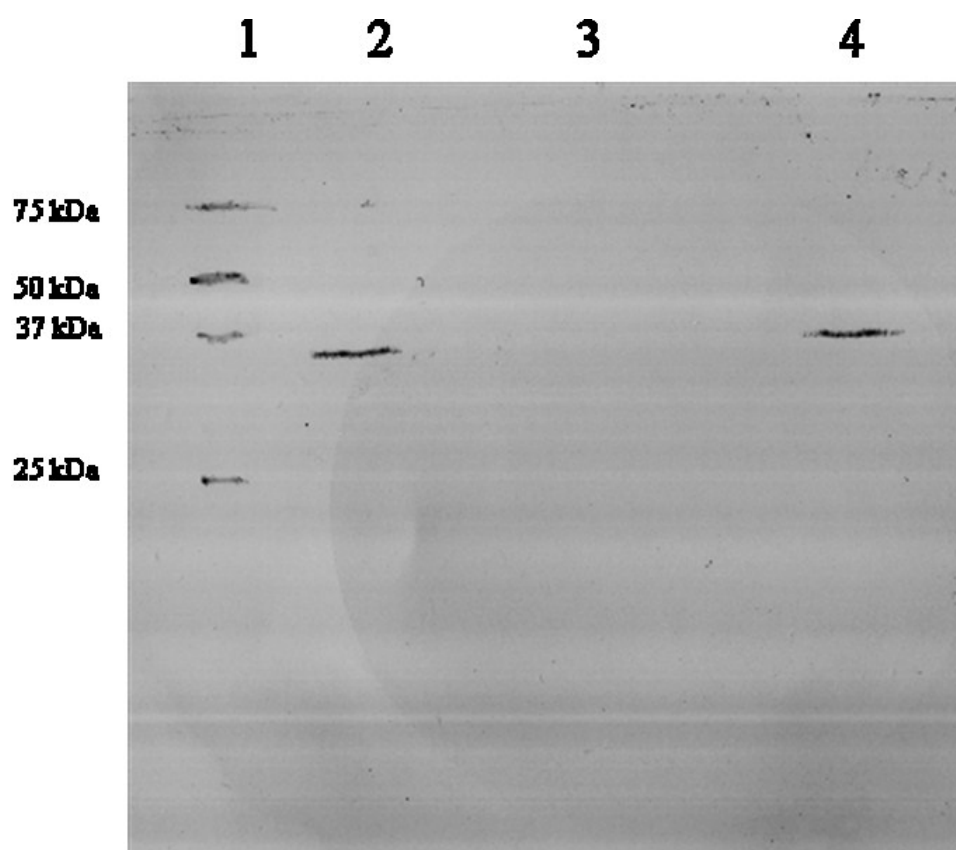


Figure 5.13: His₆-tag Western blot analyses of SeLppC and SzLppC.

Lanes were loaded as follows: Lane 1, molecular weight markers (25, 37, 50, 75 kDa); lane 2, SeLppC; lane 3, Whole cell extract of *S. equi* (negative control); lane 4, SzLppC. The purified protein amount added was 0.1 µg.

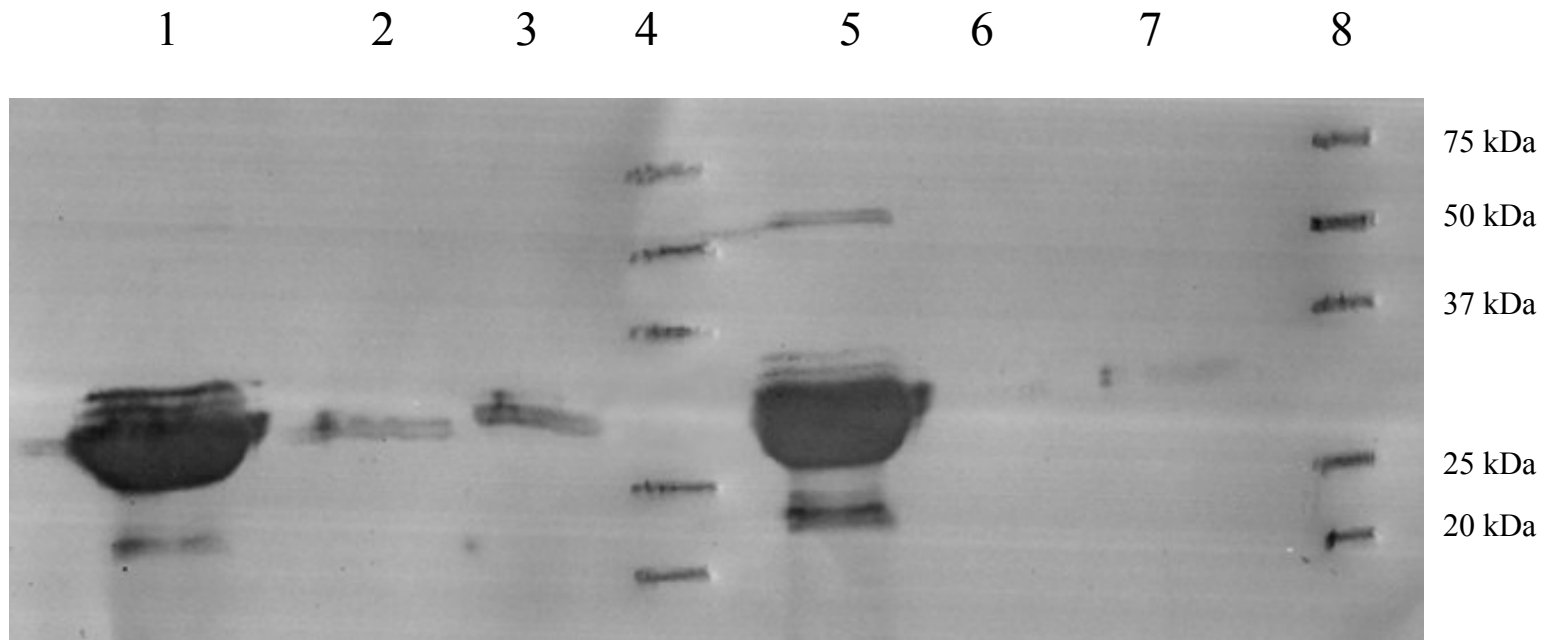


Figure 5.14: Anti-LppC Western blot analysis of purified SeLppC and SzLppC products.

The lanes were loaded as follows: Lane 1, SeLppC; lane 2, *S. equi* 4047; lane 3, *S. equi* 9672; lane 4, molecular weight markers (20, 25, 37, 50, 75 kDa); lane 5, SzLppC; lane 6, *S. zooepidemicus* H70; lane 7, *S. zooepidemicus* 3682; lane 8, molecular weight markers (20, 25, 37, 50, 75 kDa). Amount of protein added to each well was 5 μ g

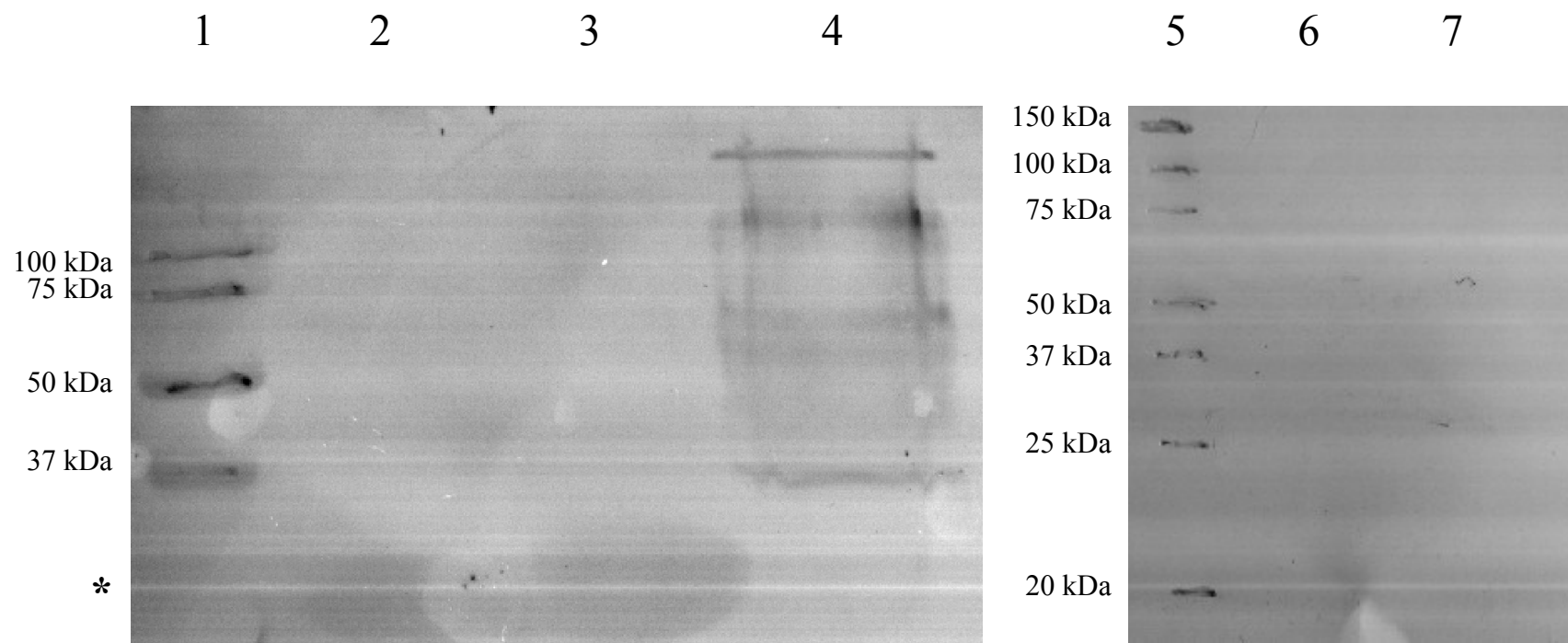


Figure 5.15: Western blot analysis of purified SeLppC and SzLppC against equine convalescent serum (pre and post infection).

The expected position of both SeLppC and SzLppC is marked with an asterisk by the left hand panel. Lanes were loaded as follows: Lane 1, molecular weight markers (37, 50, 75, 100 kDa); lane 2, SeLppC; lane 3, SzLppC; lane 4, whole cell extract of *S. equi* 4047; lane 5, molecular weight markers (20, 25, 37, 50, 75, 100, 150 kDa); lane 6, SeLppC; lane 7, SzLppC. Amount of protein added to each well was 1µg. Left hand pannel (lanes 1-4), blot probed with post-infection serum. Right hand panel (lanes 5-7), blot probed with matched pre-infection serum.

5.13. Zymographic detection of acid phosphatase activity

Purified SeLppC and SzLppC were assayed for acid phosphatase activity by zymography (see section 2.14.6). A total amount of 5 µg of purified protein was added to each well within the initial SDS-PAGE-gel. An identical SDS-PAGE replicate was also undertaken alongside zymographic analysis and stained using Coomassie blue to provide a molecular weight and protein reference point. Both purified protein samples appeared as strong bands on the zymograph, at a size consistent with the expected protein size of SeLppC and SzLppC (~32 kDa) as indicated on the Coomassie blue stained SDS-PAGE-gel replicate (Figure 5.16). These data are important in confirming the functional acid phosphatase activity of the recombinant proteins using 4-nitrophenyl phosphate (*p*NPP) as substrate.

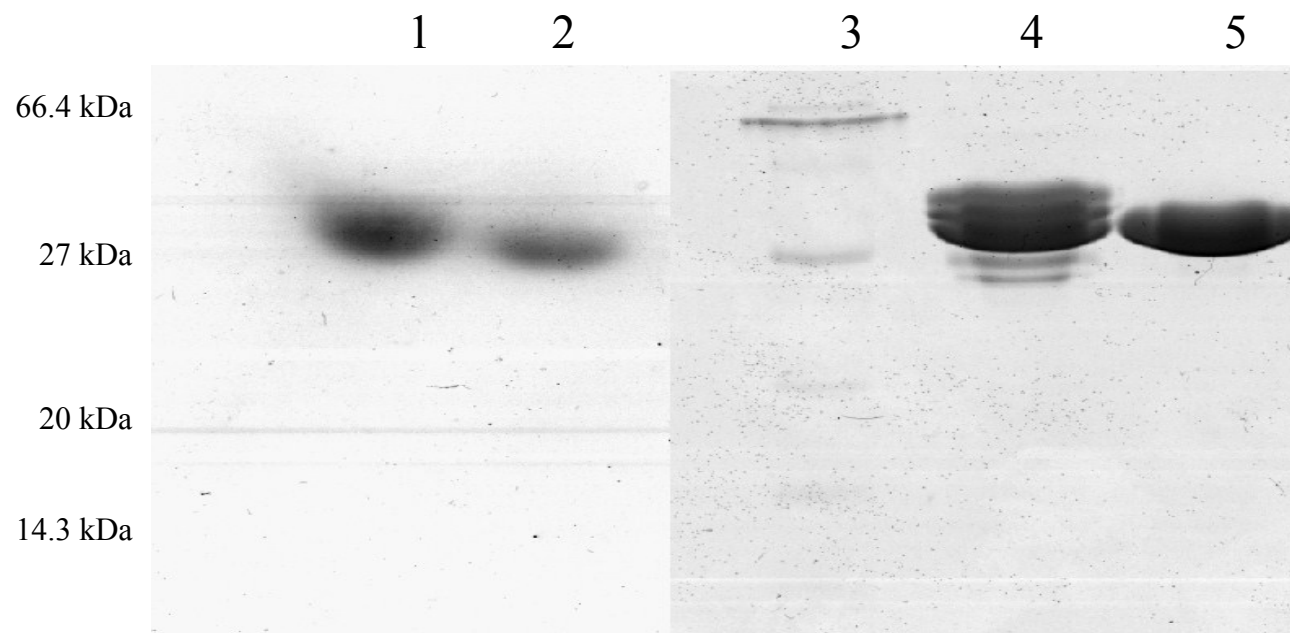


Figure 5.16: 15% SDS PAGE Zymograph of SeLppC and SzLppC and SDS PAGE (Standard Coomassie stained replicate).

Lanes were loaded as follows: Lane 1, SeLppC; lane 2, SzLppC. (These lanes are from a gel developed by zymography); lane 3, molecular weight markers (14.3, 20, 27, 66.4 kDa); lane 4, SeLppC; lane 5, SzLppC. Lanes 3-5 are from a gel developed by Coomassie blue staining. Amount of protein added to each lane was 5 μ g.

5.14. Enzymatic assays of SeLppC and SzLppC activity

5.14.1. **Michelis-Menton parameters**

Initially the activity of SeLppC and SeLppC *p*NPP was investigated. A total amount of 0.5 μ g of enzyme (determined empirically) was added to each reaction to assay SeLppC and SzLppC and the reaction was allowed to proceed for 10 minutes.

For kinetic analysis, 5 substrate concentrations (determined empirically) were chosen and reactions performed in triplicate. Linearity of each assay was confirmed prior to calculations. Prior to SeLppC and SzLppC assays a trial assay was conducted with potato acid phosphatase and the K_m calculated to be 5.77 ± 0.69 mM (Figure 5.17 and Figure 5.18).

To determine if the data reflected true Michaelis-Menten kinetics for SeLppC and SzLppC against the substrate *p*NPP, a Lineweaver-Burk plot was constructed for SeLppC and SzLppC (Figures 4.19 and 4.20 respectively). To further investigate substrate specificity, SeLppC and SzLppC were also assayed against the substrate Uridine 5' monophosphate (Figures 5.21 and 5.22 respectively) and Adenosine 5' monophosphate (Figures 5.23 and 5.24 respectively) using 3 substrate concentrations also determined empirically. Note that the acid phosphatase assay for these nucleotide monophosphate substrates cannot be measured colorimetrically and so phosphate released was determined using a phosphate standard curve (Figure 5.25).

A summary table is presented indicating the comparative K_m calculated for each substrate between SeLppC and SzLppC (Table 5.2).

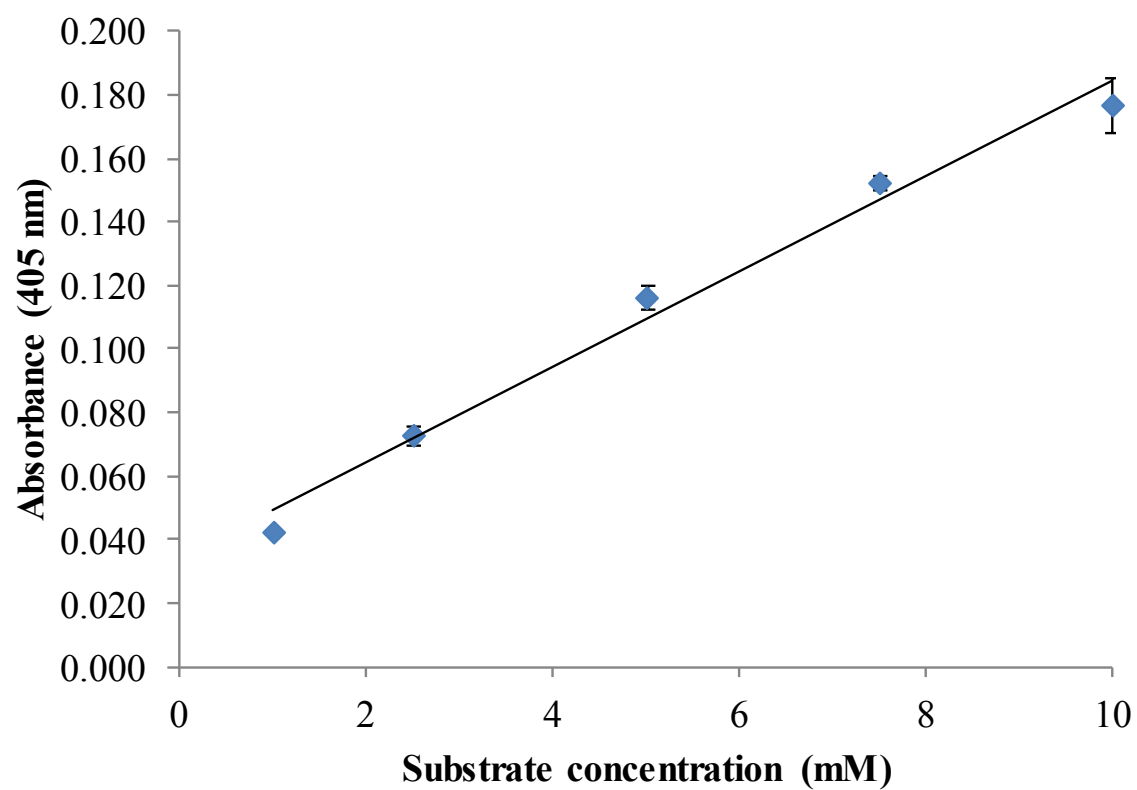


Figure 5.17: Standard absorbance of control potato acid phosphatase against NPP.

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx. 10 minutes with 5U of potato acid phosphatase.

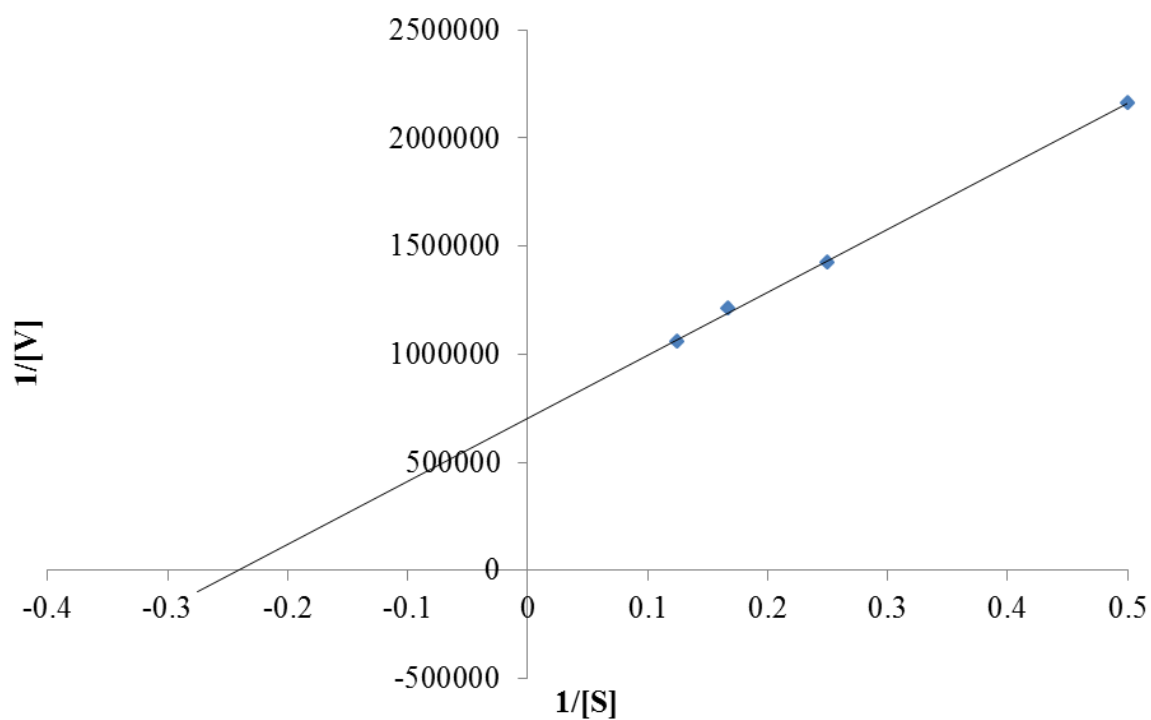


Figure 5.18: Lineweaver-Burk plot of control potato acid phosphatase against *p*NPP.

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx. 10 minutes with 5U of potato acid phosphatase.

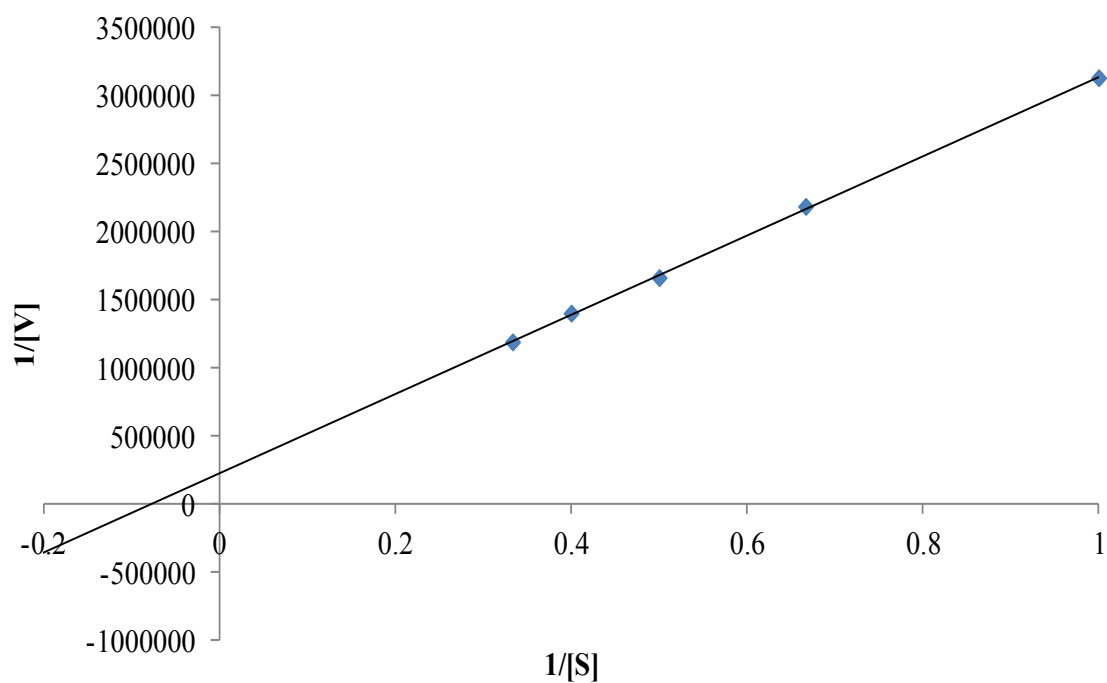


Figure 5.19: Lineweaver-Burk plot for SeLppC against *p*NPP

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx. 10 minutes with 0.5 μ g of SeLppC.

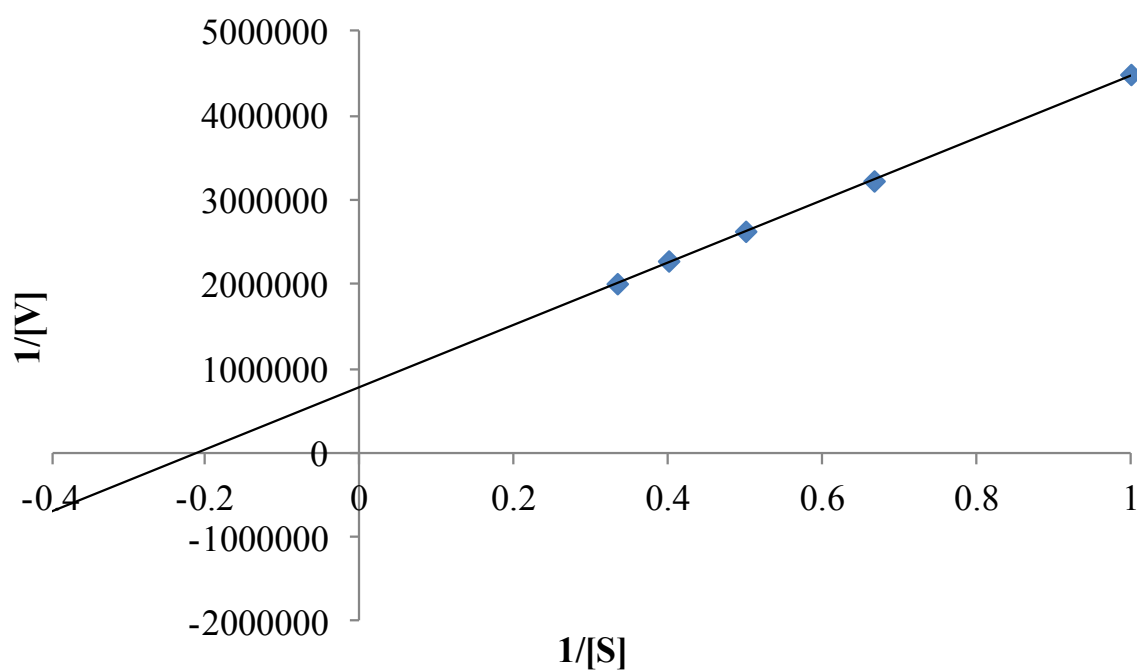


Figure 5.20: Lineweaver Burk plot for SzLppC against *p*NPP

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx 10 minutes with 0.5 μ g of SzLppC.

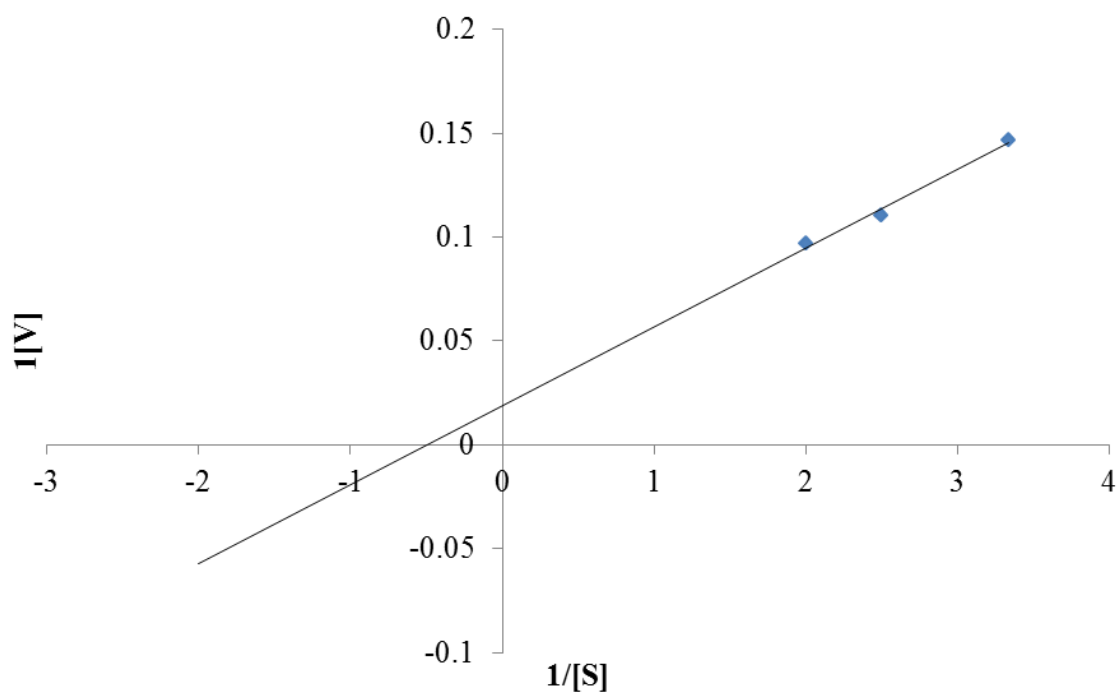


Figure 5.21: Lineweaver Burk plot for SeLppC against 5'UMP

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx. 10 minutes with 0.5 μ g of SeLppC.

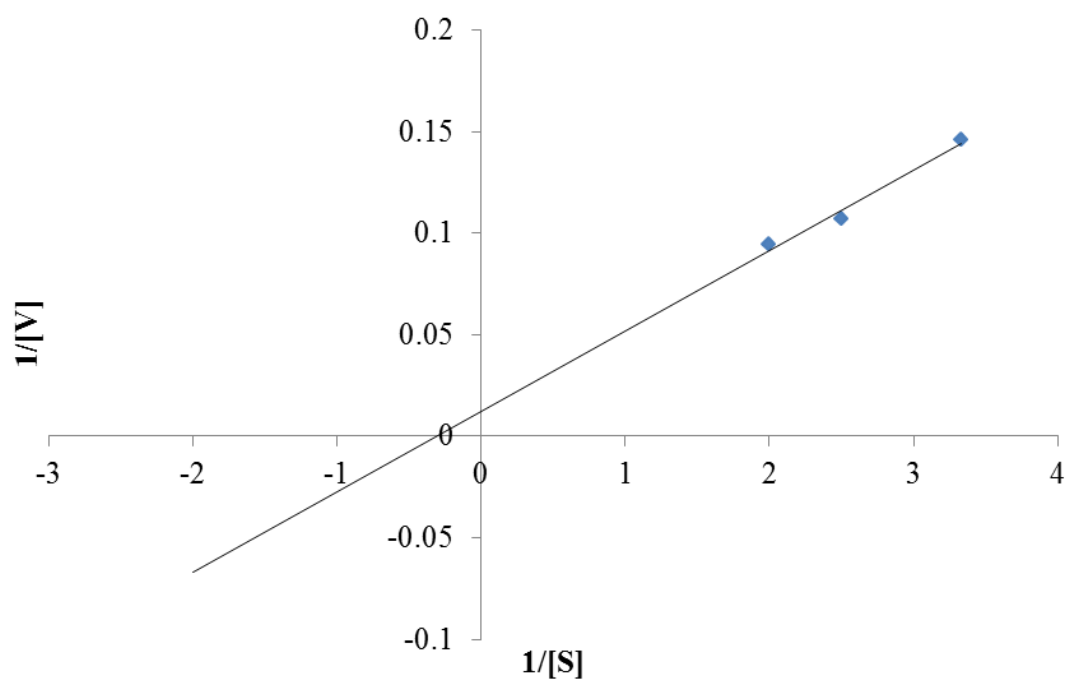


Figure 5.22: Lineweaver-Burk plot for SzLppC against 5'UMP

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx. 10 minutes with 0.5 μ g of SzLppC.

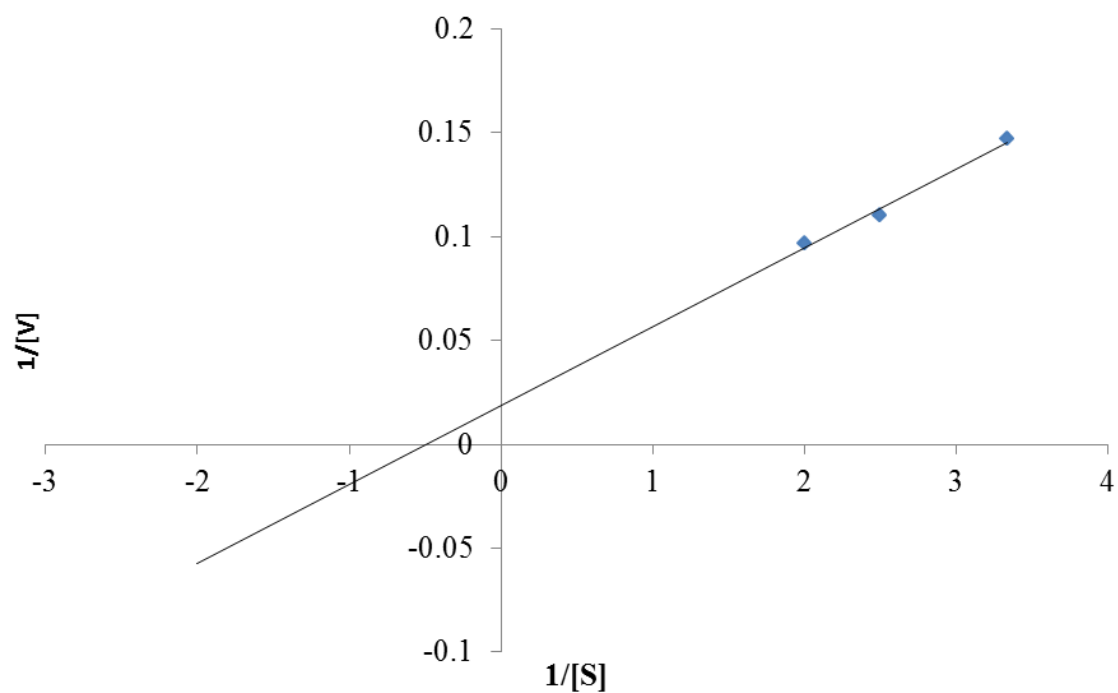


Figure 5.23: Lineweaver Burk plot of SeLppC against 5'AMP

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for approx. 10 minutes with 0.5 μ g of SeLppC.

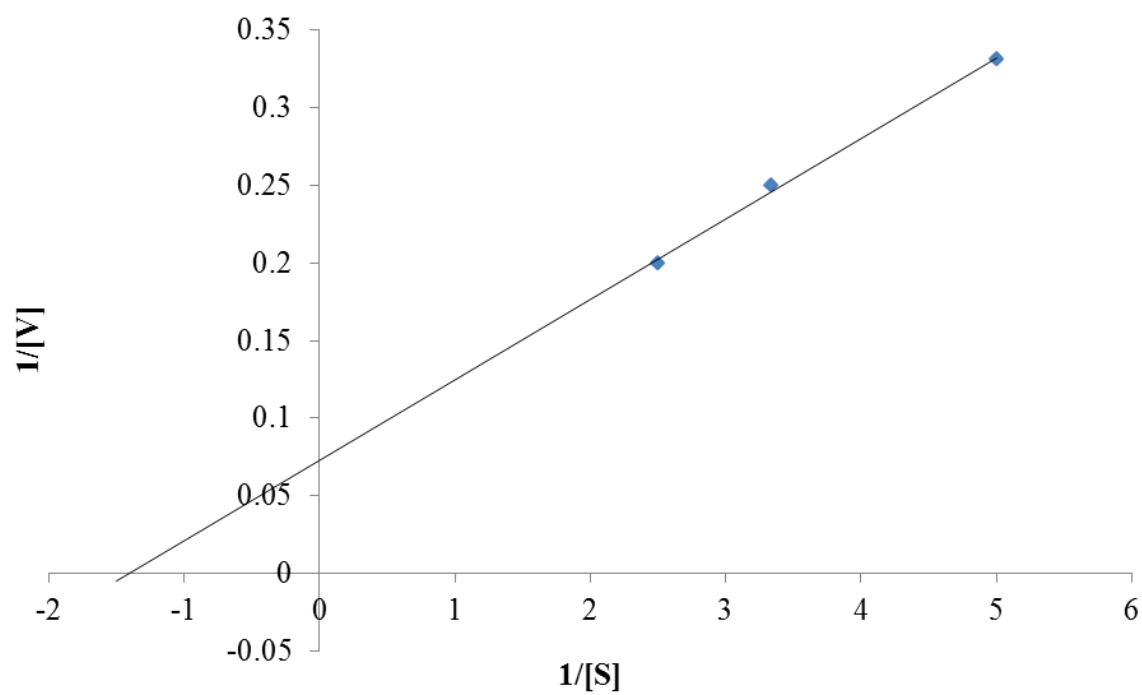


Figure 5.24: Lineweaver Burk plot of SzLppC against 5' AMP

The reactions were performed in triplicate as detailed in section 2.14.1. Each reaction was undertaken for 10 minutes with 0.5 μ g of active SzLppC.

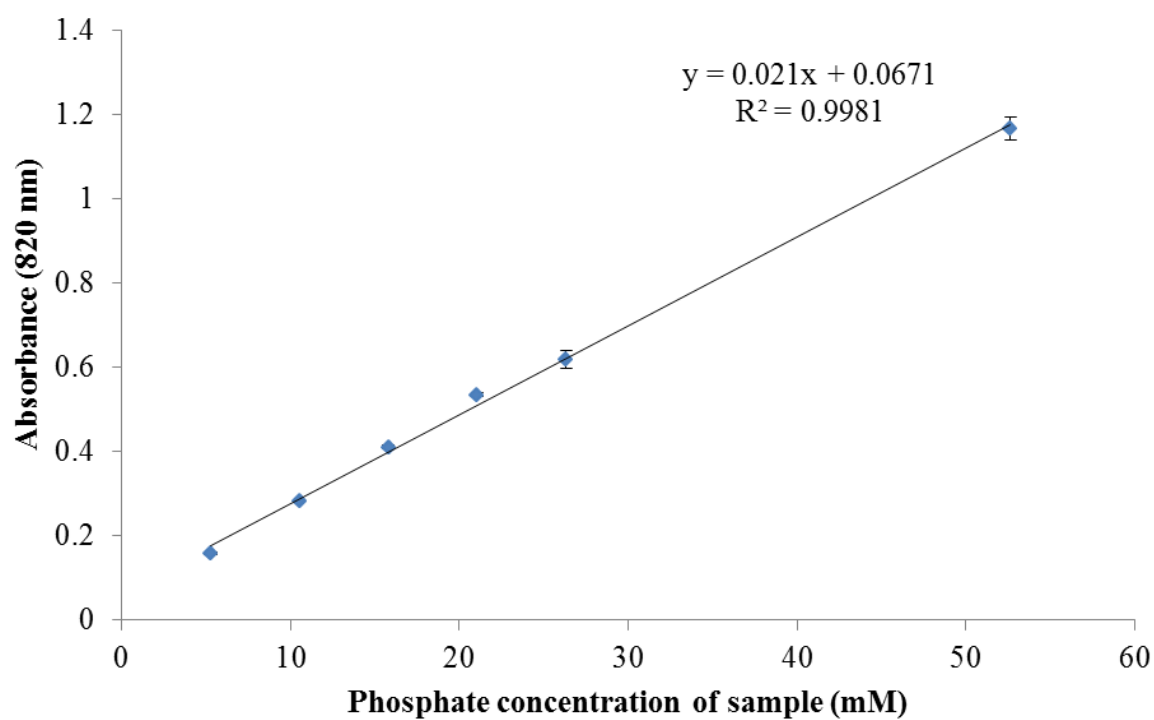


Figure 5.25: Standard curve of inorganic phosphate assay

The reactions were performed in triplicate as detailed in section 2.14.1. Error bars represent the standard deviation from the mean.

Table 5.2: Km (mM) activities of SeLppC and SzLppC

<u>Substrate</u>	<u>SeLppC</u>	<u>SzLppC</u>
<i>p</i> NPP	13.27±3.91	4.681±0.48
5' AMP	0.63±0.17	0.8152±0.18
5'UMP	1.46±0.44	2.407±1.65

*AP control (source potato) had a Km of 5.772±0.6895 mM against *p*NPP.

5.14.2. Determination of temperature optimum of SeLppC and SzLppC

An investigation was undertaken to assess the optimal temperature of enzyme activity for SeLppC and SzLppC. A total of 0.5µg of enzyme and 1mM of *p*NPP substrate was used in each reaction and the reaction was undertaken for a period of 10 minutes. Data is presented as relative activity against the standard activity of each enzyme at 37°C (Figure 5.26). SeLppC and SzLppC showed a somewhat linear increase in activity with increasing temperature. The rates of activity of both enzymes were shown to steadily increase towards the highest temperature tested (50°C), at which temperature the enzymes were shown to have a relative activity of around 120%. Standard deviations of the means remained minimal throughout absorbance readings.

5.14.3. Determination of divalent ion requirement

SeLppC and SzLppC were tested with a range of divalent ions using 0.5µg of enzyme, 1.0mM of *p*NPP under standard conditions. Data is presented as relative activity against the standard activity of each enzyme with 10 mM of Cu²⁺ added. Cu²⁺ is referred to as the reference divalent ion used in previous activity assays for SeLppC and SzLppC (Hamilton *et al.* 2000). Although the replacement of Cu²⁺ by other common divalent cations was demonstrated to result in reduced enzyme activity it was observed that the enzymes did not appear to possess a specific requirement for Cu²⁺ ions (Figures 5.27 and 5.28). Standard deviations of the means remained minimal throughout absorbance readings and both SeLppC and SzLppC produced consistent activity profiles.

5.14.4. Determination of optimum pH

The pH activity range of SeLppC and SzLppC was investigated against 5'UMP, a total of 0.5µg of each enzyme and 0.1mM of 5'UMP in each reaction. It was not possible to assess the pH optimum of SeLppC and SzLppC against *p*NPP across the full pH range as the addition of basic ions has been previously showed to interfere with the standard colorimetric acid phosphatase assay. Hamilton et al. (2000) demonstrated pH optima of 5.2 for EDTA resistant acid phosphatase in whole cell extracts of *S. equi*. An optimum pH of 5.2 is common for similar NSAP enzymes and as such data is presented as percentage relative activity to the recorded activity at pH 5.2.

As demonstrated, for both SeLppC and SzLppC, enzyme activity was maintained at a broad range between pH 4 and pH 6.5, with the rate of activity fairly consistent between the two enzymes. The highest activity was recorded at pH 4.0 with a small increase of ~10% activity detected for both enzymes. At pH 7.0 both enzymes have a significantly reduced rate of activity (~20% relative activity) indicating that both enzymes are bona fide acid phosphatases (Figure 5.29). Standard deviations of the means remained minimal throughout absorbance readings.

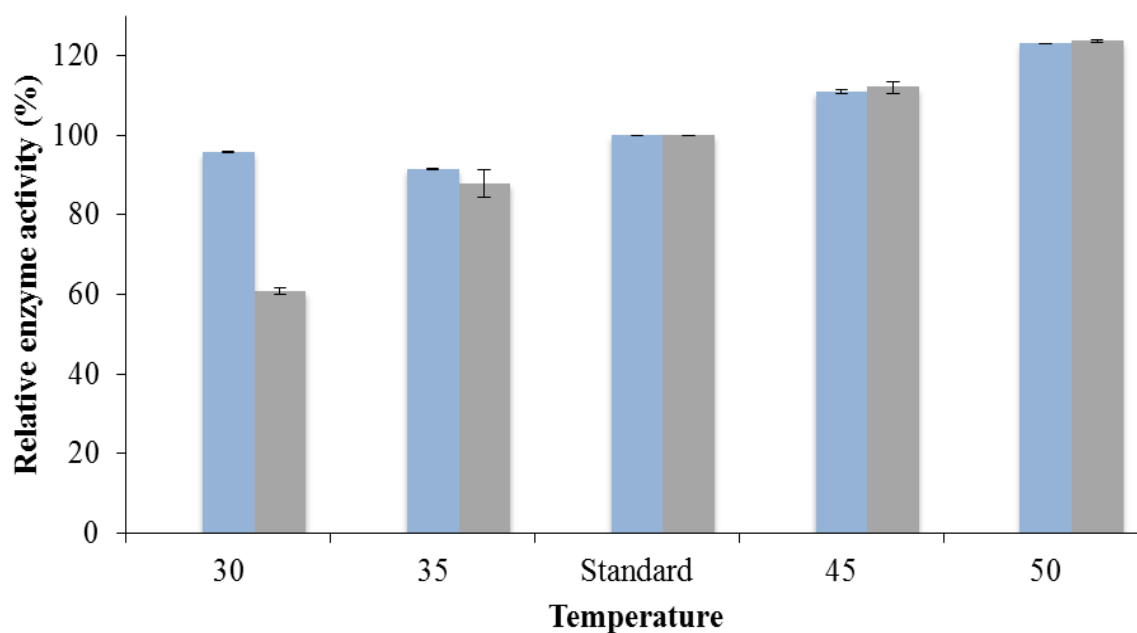


Figure 5.26: The effect of temperature on the rate of SeLppC and SzLppC enzyme activity.

The reactions were performed in triplicate as detailed in section 2.14.1. 100% activity was defined as the rate of activity of SeLppC/SzLppC at a *p*NPP concentration of 1.0 mM, utilizing 0.5µg of active SeLppC/SzLppC at pH 5.2 at a temperature of 37°C. Error bars represent the standard deviation from the mean. The blue bars represent SeLppC relative activity and the grey bars represent SzLppC relative activity

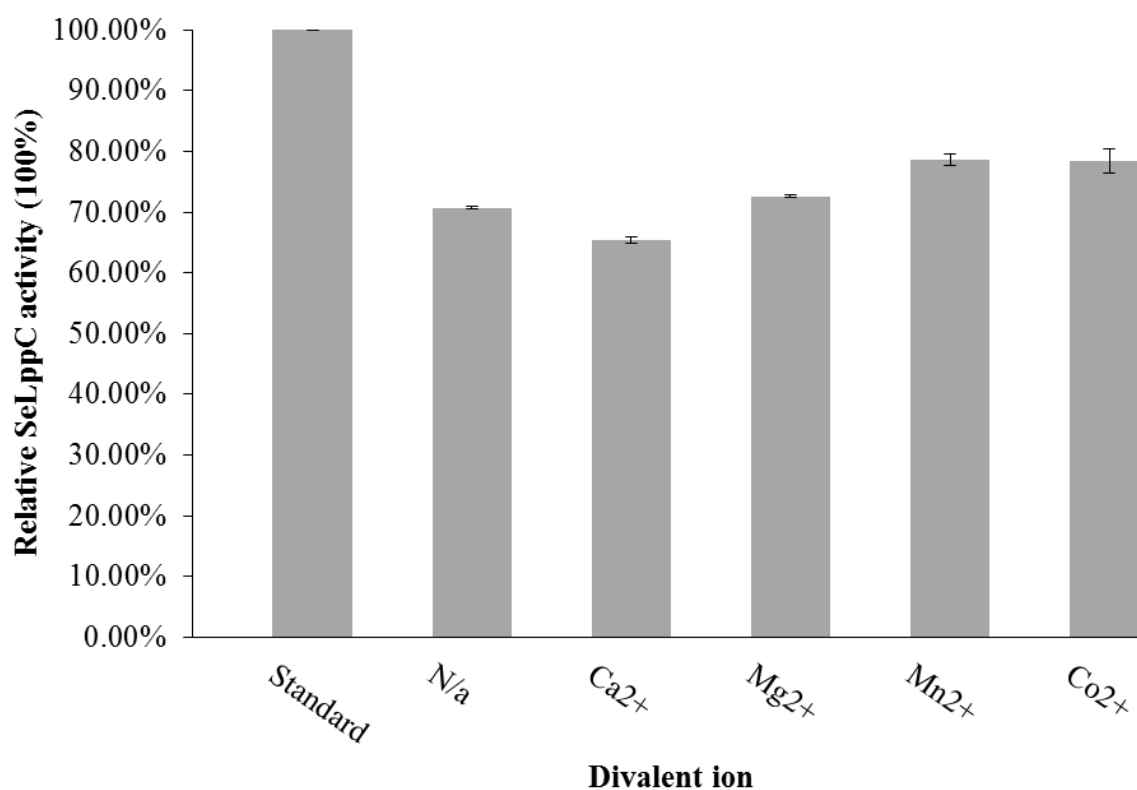


Figure 5.27: The effect of divalent cations on the rate of SeLppC enzyme activity.

The reactions were performed in triplicate as detailed in section 2.14.1. 100% activity was defined as the rate of activity of SeLppC at a *p*NPP concentration of 1.0 mM, a Cu²⁺ final concentration of 10mM and utilizing 0.5μg of active SeLppC. N/A indicates no divalent cation added. Error bars represent the standard deviation from the mean.

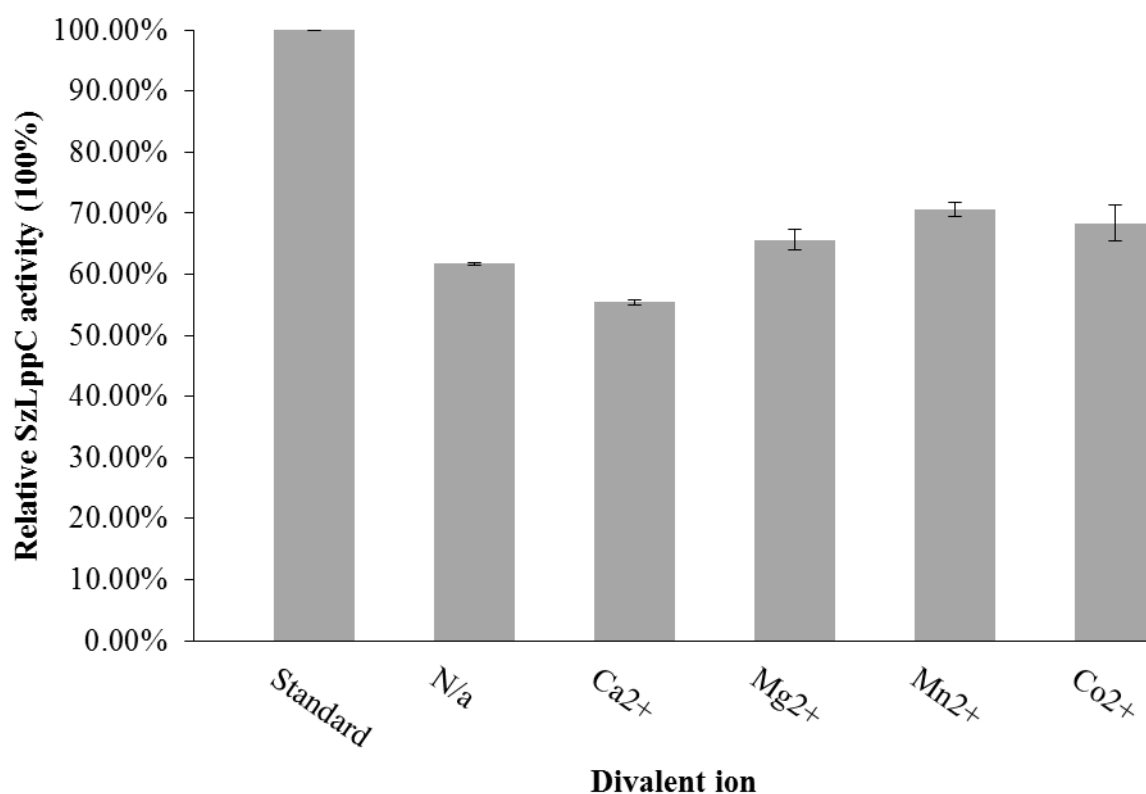


Figure 5.28: The effect of divalent cations on the rate of SzLppC enzyme activity.

The reactions were performed in triplicate as detailed in section 2.14.1. 100% activity was defined as the rate of activity of SzLppC at a *p*NPP concentration of 1.0 mM, a Cu²⁺ final concentration of 10mM and utilizing 0.5µg of active SeLppC. N/A indicates no divalent cation added. Error bars represent the standard deviation from the mean.

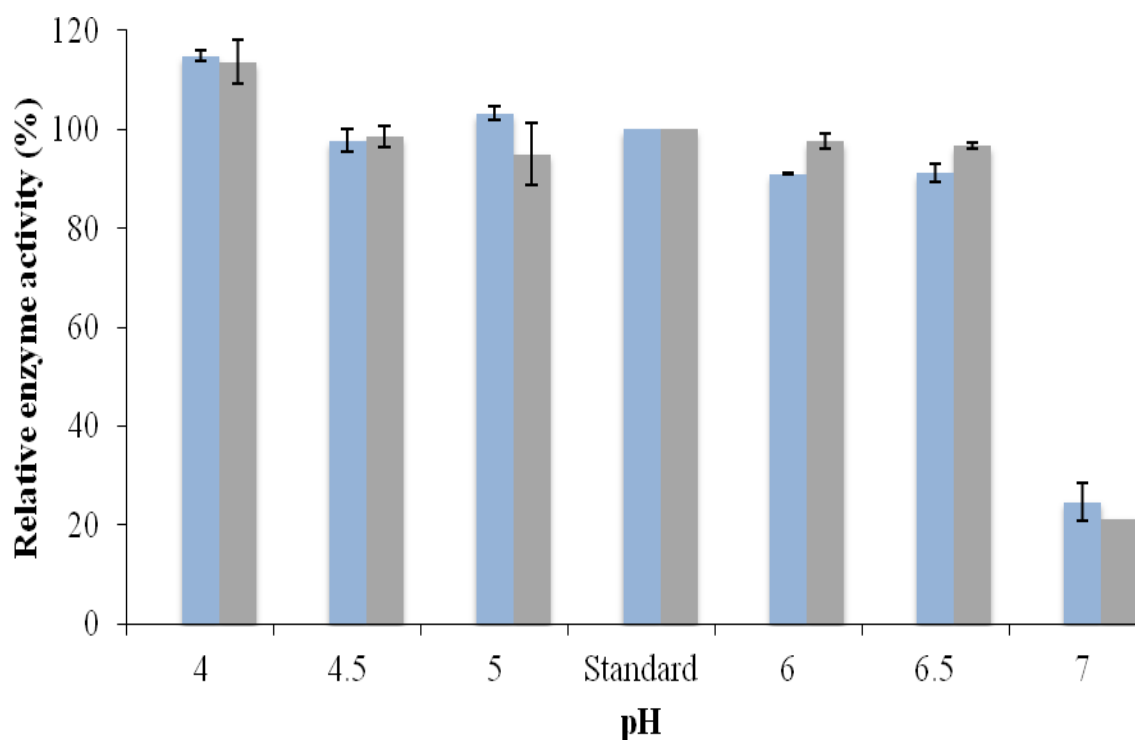


Figure 5.29: The effect of changes in pH on the rate of SeLppC and SzLppC enzyme activity.

The reactions were performed in triplicate as detailed in section 2.14.1. The reference 100% activity was defined as the rate of activity of SeLppC/SzLppC at a 5'UMP concentration of 0.1 mM, utilizing 0.5 μ g of active SeLppC/SzLppC at pH 5.2. The blue bars represent SeLppC relative activity and the grey bars represent SzLppC relative activity.

5.15. Discussion and future work.

Bioinformatic analysis of SeLppC indicated that as expected a close homologue is also encoded in the genome of *S. zooepidemicus* H70, SzLppC. Comparing the sequences of *SeLppC* and *SzLppC* obtained from the genomes of *S. equi* 4047 and *S. zooepidemicus* SzH70 indicated a deletion of one base between the nucleotide position 1887168 and 1887169 of the *S. zooepidemicus* H70 genome. However sequence analysis conducted in this study indicated that the sequence deposited from the *S. zooepidemicus* H70 genome project was incorrect i.e. indicating that this strain does have the capacity to synthesise a full length LppC (as with the *S. zooepidemicus* MGCS10565 genome).

Western blot analysis of SeLppC and SzLppC using LppC specific antisera indicated that antibody detection was significantly poorer in *S. zooepidemicus* compared to *S. equi* whole cell extracts. The correction of the single base pair deletion in the signal peptide coding region of *szo18670* demonstrates that this is not because *szo18670* is a pseudogene.

As shown in this study, the sequences encoding SeLppC and SzLppC were successfully cloned into pLysE (DE3) BL21 *E. coli* for protein expression. The use of pLysE has been previously shown to produce high levels of the closely related NSAP, e(p4) (Reilly 1999). Both *SeLppC* and *SzLppC* were demonstrated to be present in the presumptive pSeLppC and pSzLppC constructs by restriction enzyme digestion and short confirmatory sequencing. Reasonable quantities of purified SeLppC and SzLppC proteins were also successfully obtained with post-FPLC, with SDS-PAGE analysis indicating a high level of purification. The presence of a histidine N-terminal tag was also verified by Western blotting. Zymography also confirmed that both SeLppC and SzLppC were acid phosphatases.

Furthermore Western blot analysis of the recombinant SzLppC indicated that discrepant molecular weight (Figure 5.7) between SzLppC from strain H70 and SeLppC (e.g. from strain 4047) was limited to whole cell extracts of SzLppC, suggesting perhaps some specific proteolytic processing. Failure to detect SzLppC in whole cell extracts of other strains could perhaps be explained by a discrepancy between levels of expression of *SzLppC* and *SeLppC* in vitro.

Further Western blot analysis of SeLppC and SzLppC using equine sera obtained post *S. equi* infection indicated a very weak detection for both SeLppC and SzLppC. This may indeed be consistent with SeLppC and SzLppC expression in vivo and suggests that both enzymes are only weakly immunogenic.

As previously described, Class C NSAP have fairly broad substrate specificity and as expected the arylphosphate *p*-NPP represented a readily hydrolysed substrate for SeLppC and SzLppC with recorded K_m values of 13.27 ± 3.91 and 4.68 ± 0.482 mM respectively. Of greater significance was the hydrolysis of the 5' nucleoside monophosphates, AMP and UMP by SeLppC and SzLppC. Both enzymes were shown to possess a greater affinity to 5'UMP than 5'AMP.

It has not been determined whether or not either SeLppC or SzLppC have affinity towards 3' nucleoside monophosphates. The closely related Class C NSAP, e (P4) of *H. influenzae* (79/229 amino acid sequence identity to SeLppC) has previously been shown to have little affinity to 3' nucleoside monophosphates with a suggestion that e(p4) of *H. influenzae* discriminates between the 3' and 5' phosphomonoester binding within the active site (Reilly and Calcutt 2004). Enzyme kinetic comparisons of SeLppC/SzLppC with other similar enzymes such as e(P4) of *H. influenzae* and LppC of *S. equisimilis* are difficult

because neither of these enzymes were subjected to detailed kinetic analysis. Comparisons can however be made with the enzymes OlpA (Passariello 2003), rCppa (Reilly 2009) and rHppa (Reilly & Calcutt, 2004) (Table 5.3). SeLppC and SzLppC were demonstrated to have a far higher K_m to *p*NPP than OlpA. Interestingly SeLppC and SzLppC appear to have substantially higher K_m to 5'UMP than any of the other similar enzymes previously analysed. Throughout kinetic enzyme studies of SeLppC and SzLppC, SzLppC was demonstrated to have higher K_m values for all substrates compared to SeLppC, although both enzymes possessed comparative kinetic trends. Since the sequencing data (Figure 5.5) has demonstrated that the mature SeLppC and SzLppC are nearly identical in terms of amino acid sequence it is unlikely that enzyme kinetic analysis variations are a result of genuine biochemical differences. A likely explanation of this discrepancy is that the obtained purified proteins have differing levels of purification.

The activity of several class C enzymes has been shown to possess a requirement for divalent cations. As demonstrated in this study, neither SeLppC nor SzLppC possessed an absolute requirement for divalent cations; both enzymes retained a strong catalytic activity in apparent absence of any divalent ion. These findings are consistent with studies of the e (P4) acid phosphatase, in which some activity was retained after dialysis against buffer containing EDTA followed by dialysis against buffer to remove the chelator (Reilly 1999).

It was also demonstrated that SeLppC and SzLppC possess a selective affinity for divalent cations with only the presence of Cu^{2+} demonstrated to increase catalytic activity amongst a group of the most commonly known divalent cations affecting the reported activities of acid phosphatases. These findings place SeLppC and SzLppC in a similar group as e (P4) (Reilly *et al.* 1999; Reilly *et al.* 2001), HppA (Reilly *et al.* 2006) and LppC (Malke

1998) with a demonstrated affinity towards Cu^{2+} as opposed to OlpA (Ou *et al.* 2006) and SapS (Du Plessis *et al.* 2002) which possess a demonstrated preference towards Mg^{2+} .

As expected, both SeLppC and SzLppC had a demonstrated optimal pH activity at pH 5 or less, an optimum shared with several members of its class. However a strong level of activity was retained across a fairly broad range between pH 4.0 and pH 6.5. This is contrasting to the demonstrated pH optima of LppC of *S. equisimilis* which featured a sharp peak at pH ~5 with significantly reduced activity demonstrated between pH 3-4.5 and pH 5.5-6.5 (Malke 1998). It has yet to be determined if this difference in pH range has biological significance *in vivo*. The effect of pH could not be determined accurately for SeLppC and SzLppC against pNPP.

The results of the investigation of temperature against the activities of SeLppC and SzLppC indicated a high level of enzyme activity was maintained at temperatures of at least 50°C. Although unexpected, this may be consistent with the data reported here for SlaA (see chapter 4), in that an enzyme's ability to remain potent enzymically at higher temperatures is advantageous during the strangles disease state in which an infected pyrexia horse can exhibit elevated temperatures above (>40°C).

A well-documented characteristic of all class C NSAP enzymes is a DDDD motif assumed to be of catalytic significance. This has been experimentally verified in both CppA of *C. perfringens* (Reilly *et al.* 2009) and e (P4) of *H. influenzae* but it remains to be determined whether or not this is also the case for SeLppC or SzLppC, although this would be expected to be so. By undertaking site directed mutagenesis of one or more of the aspartic

acid residue of this motif, it would be expected that mutant SeLppC and SzLppC would also exhibit reduced zymogram detection and a significantly reduced enzyme activity.

A further area of interest regards the level of conservation of SeLppC and SzLppC amongst different *S. equi* and *S. zooepidemicus* strains. Since the activities of NSAP have been known to vary amongst closely related strains it would be interesting to reveal whether or not different strains possess different enzyme profiles. The strong reaction of a range of whole cell extracts of *S. equi* to LppC specific antisera indicates that it is probable that SeLppC is consistently well expressed amongst *S. equi* strains whereas lower/anomalous cross reactions were obtained with *S. zooepidemicus* strains. However, a more extensive strain collection must be investigated under highly standardised growth and protein preparation conditions. Moreover it may be possible that short sequencing of respective coding sequences amongst a variety of strains of the two subspecies may reveal differences. Alternatively, the availability of further whole genome sequences of each species would allow for this.

It remains to be determined whether or not SeLppC and SzLppC play a critical role in the survival of either or both bacteria *in vivo* or *in vitro*. Mutagenesis studies need to be carried out to address this. The demonstrated hydrolysis of 5' nucleoside monophosphates indicates that SeLppC and SzLppC are likely to constitute a nucleotide biosynthetic salvage pathway. This conclusion is consistent with the role proposed to LppC of *S. equisimilis* and CppA of *C. perfringens* (Reilly *et al.* 2009).

The production of an isogenic mutant strain of *SeLppC* in *S. equi* would provide the tools to elucidate the role of SeLppC in virulence. It would also be interesting whether or not a Δ *SeLppC* strain was defective in intramacrophage survival/replication, similar to that demonstrated by Δ *acpA* of *Francisella* spp (Mohapatra 2007). Furthermore an obvious

further experiment involving a $\Delta SeLppC$ strain would be the investigation of possible reduced virulence in either the mouse or equine models of infection.

Table 5.3: Recorded K_m 's of SeLppC and SzLppC for three substrates and other closely related NSPC's

Data obtained in this study is highlighted in bold.

<u>Enzyme</u>	<u>Substrate</u>	<u>K_m (mM)</u>	<u>Reference</u>
SeLppC	4-nitrophenyl phosphate	13.27±3.910	N/a
SzLppC	4-nitrophenyl phosphate	4.681±0.4820	N/a
OlpA	4-nitrophenyl phosphate	0.180 ± 0.06	(Passariello, et al. 2003)
SeLppC	5' AMP	0.23±0.04	N/a
SzLppC	5'AMP	0.8152±0.1789	N/a
rCppA	5'AMP	0.269±0.05	(Reilly, et al. 2009)
rHppA	5'AMP	1.06± 0.18	(Reilly and Calcutt 2004)
OlpA	5'AMP	0.016±0.001	(Passariello, et al. 2003)
SeLppC	5'UMP	1.459±0.4386	N/a
SzLppC	5'UMP	2.407±1.651	N/a
rCppA	5'UMP	0.221±0.03	(Reilly, et al. 2009)
rHppA	5'UMP	0.91±0.03	(Reilly and Calcutt 2004)
sOlpA	5' UMP	0.013±0.08	(Passariello, et al. 2003)

6. Conclusion

As discussed, the likelihood of an improved vaccine against equine strangles relies on the identification and characterisation of novel immunogens followed by the development of combinatorial subunit vaccines based on those immunogens deemed to provide sufficient protection in equine challenge studies (Timoney *et al.* 2007, Florindo *et al.* 2009, Waller and Jolly 2007b). Recent advancements towards a recombinant vaccine have clearly been hastened by the availability of the genomes of *S. equi* and *S. zooepidemicus* thus allowing the identification of many novel surface exposed and secreted proteins

Herein this thesis we have presented the investigation of three novel pairs of proteins; the covalently attached, surface exposed proteins SEQ2190 and SEQ2180, the secreted phospholipase A₂ enzymes SlaA and SlaB and the lipoprotein encoded Acid phosphatase enzymes SeLppC and SzLppC. All of the afore-mentioned proteins were initially identified following genome comparisons of the *S. equi* and *S. zooepidemicus* of genomes and demonstrated to possess uniqueness in relation to *S. equi*.

SEQ2190 was demonstrated to be unique in *S. equi* (with a *seq2190* paralogue pseudogene in *S. zooepidemicus*) and had previously been shown to be immunogenic. This cell surface associated protein was also demonstrated in this study to possess a close homologue (SEQ2180) with a similar paralogue in *S. zooepidemicus* (SZO18890). Bioinformatic structural analysis of SEQ2190 conducted in this study has also revealed a putative structure dominated by an alpha helical stalk protruding through the hyaluronic acid capsule, with highly probable putative host ligand interaction.

Although the exact function(s) of SEQ2190 and SEQ2180 in relation to virulence remains unknown and three dimensional structures of either protein were not successfully solved in this study during the time-frame given, constructs with the ability to express both proteins with N-terminal Histidine tags were successfully created as demonstrated. An N-terminal portion of SEQ2190 was also shown to be wholly unique to *S. equi* and a construct was also created in this study expressing only this unique, truncated protein.

The secreted phospholipase A₂, SlaA and SlaB was also investigated in this thesis. SlaA was demonstrated to be unique to *S. equi* and bear close similarity to the investigated virulence factor SpSlaA of *S. pyogenes*. Although some studies had previously been undertaken (Nagiec *et al* 2004) this study has further characterised SlaA utilising constructs expressing N-terminal Histidine tagged SlaA.

Using the created constructs it has been possible to obtain purified SlaA and gain further insight into the enzyme kinetics of this enzyme. The data obtained from this study has now shown that SlaA is active against DHT-PC but not AT-PC, has a optimum temperature of 45°C and a specific requirement for Ca²⁺. The K_m and specific activity values for SlaA have also now been recorded. Western blot analysis has also indicated that antibodies are indeed raised against SlaA in the early onset of a strangles infection.

Finally, the acid phosphatases SeLppC and SzLppC were investigated. Although SeLppC and SzLppC had been previously investigated and characterised (Hamilton *et al.* 2000), characterisation was carried out using only whole cell extracts of *S. equi* and *S. zooepidemicus* strains (non-genome strains) and undertaken prior to the availability of the *S. equi* and *S. zooepidemicus* genomes. Additionally it was thought that the genome strain possessed a ‘frame-shift’ mutation, rendering SeLppC a pseudogene.

Following this study it can now be concluded that SeLppC in *S. equi* H70 is not a pseudogene. Furthermore it has been shown that SeLppC and its close paralogue SzLppC are active against the substrate *p*NPP and the biologically relevant substrates 5'-AMP and 5'UMP, possessing activity at temperatures up to at least 50°C, possessing the strongest activity at pH <5 and possessing no apparent specific requirement to Cu²⁺.

Finally, it is hoped that the data generated during this study will be informative towards the understanding of the molecular basis of virulence of *S. equi* and that the constructs created will be further used to investigate these potential virulence factors further in view of the creation of mutant strains of *S. equi* suitable for equine challenge studies in the future.

References

- Akporiaye, E. T. and Baca, O. G. (1983) Superoxide anion production and superoxide dismutase and catalase activities in *Coxiella burnetii*. *Journal of Bacteriology*, 154(1):520-523.
- Akporiaye, E. T., Rowatt, J. D., Aragon, A. A., Baca, O.G. (1983) Lysosomal response of a murine macrophage-like cell-line persistently infected with *Coxiella burnetii*. *Infection and Immunity*, 40(3):115-1162.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215(2):403-10.
- Andersson, H., Bakker, E. & G. Vonheijne, G. (1992) Different Positively charged amino-acids have similar effects on the topology of a polytopic transmembrane protein in *Escherichia Coli*. *Journal of Biological Chemistry*, 267(3):1491-1495.
- Anzai, T., Kuwamoto, Y., Wana, R., Sugita, S., Kakuda, T., Takai, S., Higuchi, T., Timoney, J. F. (2005) Variation in the N-terminal region of an M-like protein of *Streptococcus equi* and evaluation of its potential as a tool in epidemiologic studies. *American Journal of Veterinary Research*, 66(12):2167-2171.
- Anzai, T., Timoney, J. F., Kuwamoto, Y., Fujita, Y., Wada, R., Inoue, T. (1999) In vivo pathogenicity and resistance to phagocytosis of *Streptococcus equi* strains with different levels of capsule expression. *Veterinary Microbiology*, 67(4):277-286.
- Artiushin, S. C., Timoney, J. F., Sheoran, A. S., Muthupalani, S. K. (2002) Characterization and immunogenicity of pyrogenic mitogens SePE-H and SePE-I of *Streptococcus equi*. *Microbial Pathogenesis*, 32(2):71-85.

- Asanuma, M., Kurokawa, K., Ichikawa, R., Kyong-Hwa, Ryu, K., Chae, J., Dohmae, N., Lee, B. L., Nakayama, H. (2011) Structural evidence of alpha-aminoacylated lipoproteins of *Staphylococcus aureus*. *Febs Journal*, 278(5):716-728.
- Baca, O. G., Baca, O. G., Roman, M. J., Glew, R. H., Christner, R. F., Buhler, J. E., Aragon, A. S (1993) Acid-phosphatase-activity in *Coxiella Burnetii* - A possible virulence factor. *Infection and Immunity*, 61(10):4232-4239.
- Baiter, S., Benin, A., Pinto, S. W. L., Teixeira, L. M., Alvim, G. G., Luna, E., Jackson, D., LaClaire, L., Elliott, J., Facklam, R., Schuchat, A. (2000) Epidemic nephritis in Nova Serrana, Brazil. *The Lancet*, 355(9217):1776-1780.
- Bagos, P. G., Tsirigos, K. D., Liakopoulos, T. D., Hamodrakas, S. J. (2008) Prediction of lipoprotein signal peptides in Gram-positive bacteria with a Hidden Markov Model. *Journal of Proteome research*, 7(12):5082-93.
- Banks, D. J., Lei, B. F and Musser, J. M (2003) Prophage induction and expression of prophage-encoded virulence factors in group A Streptococcus serotype M3 strain MGAS315. *Infection and Immunity*, 71(12):7079-7086.
- Baptiste, K. E. (1998) A preliminary study on the role of the equine guttural pouches in selective brain cooling. *Veterinary Journal*, 155(2):139-148.
- Baptiste, K. E., Naylor, J. M., Bailey, J., Barber, E. M., Post, K., Thornhill, J. (2000) Physiology - A function for guttural pouches in the horse. *Nature*, 403(6768):382-383.
- Barinov, A., Loux, V., Hammani, A., Nicolas, P., Langella, P., Ehrlich, D., Maguin, Van de Guchte, M. (2009) Prediction of surface exposed proteins in *Streptococcus pyogenes*, with a potential application to other Gram-positive bacteria. *Proteomics*, 9(1):61-73.

Båverud, V., Johansson, S. K and Aspan, A (2007) Real-time PCR for detection and differentiation of *Streptococcus equi* subsp. *equi* and *Streptococcus equi* subsp. *zooepidemicus*. *Veterinary Microbiology*, 124(3-4):219-229.

Bentley, S. D., Chater, K.F., Cerdeno-Tarrage, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabinowitz, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., Hopwood, D. A (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*, 417(6885):141-147.

Beres, S. B., Sylva, G. L., Barbian, K. D., Lei, B. F., Hoff, J. S., Mammarella, N. D., Liu, M. Y., Smoot, J. C., Porcella, S. F., Parkins, L. D., Campbell, D. S., Smith, T. M., McCormick, J. K., Leung, D. Y. M., Schlievert, P. M. (2002) Genome sequence of a serotype M3 strain of group A *Streptococcus*: Phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proceedings of the National Academy of Sciences of the United States of America*, 99(15):10078-10083.

Bieker, K. L., Phillips, G. J and Silhavy, T. J. (1990) The Sec and Prl genes of *Escherichia coli*. *Journal of Bioenergetics and Biomembranes*, 22(3):291-310.

Bisno, A. L., Collins, C. M and Turner, J. C (1997) M proteins of group C streptococci isolated from patients with acute pharyngitis. *Journal of Clinical Microbiology*, 35(1):338-338.

- Blank, L. M., Hugenholtz, P and Nielsen, L. K. (2008) Evolution of the hyaluronic acid synthesis (has) operon in *Streptococcus zooepidemicus* and other pathogenic streptococci. *Journal of Molecular Evolution*, 67(1):13-22.
- Blank, L. M., McLaughlin, R. L and Nielsen, L. K. (2005) Stable production of hyaluronic acid in *Streptococcus zooepidemicus* chemostats operated at high dilution rate. *Biotechnology and Bioengineering*, 90(6):685-693.
- Boschwitz, J. S. and Timoney, J. F. (1994) Characterization of the antiphagocytic activity of equine fibrinogen for *Streptococcus equi* subsp *equi*. *Microbial Pathogenesis*, 17(2):121-129.
- Boschwitz, J. S. and Tomoney, J. F. (1994) Inhibition of C3 deposition on *Streptococcus equi* subsp *equi* by M-protein - A mechanism for survival in equine blood. *Infection and Immunity*, 62(8):3515-3520.
- Braun, M. A., Gerlach, D., Hartwig, U. F., Ozegowski, J. H., Romagne, F., Carrel, S., Kohler, W., Fleischer, B. (1993) Stimulation of Human T-Cells by Streptococcal superantigen erythrogenic toxins (Scarlet fever toxins). *Journal of Immunology*, 150(6):2457-2466.
- Braun, V., Wu, H. C., Ghuysen, J. M., Hakenbeck, R. (1994). Chapter 14 Lipoproteins, structure, function, biosynthesis and model for protein export. *In* New Comprehensive Biochemistry. Pp. 319-341: Elsevier.
- Briles, D. E., Tart, R. C., Swiatlo, E., Dillard, J. P., Smith, P., Benton, K. A., Ralph, B. A., Brooks-Walter, A., Crain, M. J., Hollingshead, S. K., McDaniel, L. S. (1998) Pneumococcal diversity: Considerations for new vaccine strategies with emphasis on pneumococcal surface protein A (PspA). *Clinical Microbiology Reviews*, 11(4):645.

- Bryant, A. E. and Stevens, D. L. (1996) Phospholipase C and perfringolysin O from *Clostridium perfringens* upregulate endothelial cell-leukocyte adherence molecule 1 and intercellular leukocyte adherence molecule 1 expression and induce interleukin-8 synthesis in cultured human umbilical vein endothelial cells. *Infection and Immunity*, 64(1):358-362.
- Bryant, S., Brown, K. K., Lewis, S., Stewart, R. C., Parizek, R. (1985). Protection against strangles with an enzymtic *Streptococcus equi* extract. *Veterinary Medicine*, 80(9):58-&.
- Buist, G., Ridder, A. N. J. A., Kok, J., Kuipers, O. P. (2006) Different subcellular locations of secretome components of Gram-positive bacteria. *Microbiology-Sgm*, 152:2867-2874.
- Burke, J. E. and Dennis, D. A (2009) Phospholipase A(2) Biochemistry. *Cardiovascular Drugs and Therapy*, 23(1):49-59.
- Cabanes, D., Dehoux, P., Dussurget, O., Frangeul, L., Cossart, P. (2002) Surface proteins and the pathogenic potential of *Listeria monocytogenes*. *Trends in Microbiology*, 10(5):238.
- Cabral, G. A. (2005) Lipids as bioeffectors in the immune system. *Life sciences*, 77(14):1699-1710.
- Camilli, A., Tilney, L. G. and Portnoy, D. A. (1993) Dual roles of PLCA in *Listeria monocytogenes* pathogenesis. *Molecular Microbiology*, 8(1):143-157.
- Center for Disease Control (1983) Group C streptococcal infections associated with eating homemade cheese, New mexico. *Morbidity and Mortality Weekly Report*, 32:510-16.
- Chanter, N., Colins, N., Holmes, N., Binns, M., Mumford, J. (1997) Characterization of the Lancefield group C *Streptococcus* 16S-23S RNA gene intergenic spacer and its potential for identification and sub-specific typing. *Epidemiology and Infection*, 118(2):125-135.

- Chanter, N., Smith, K. C. and Mumford, J. A. (1995) Equine strangles modeled in mice. *Veterinary Microbiology*, 43(2-3):209-218.
- Chanter, N., Talbot, N. C., Newton, J. R., Hewson, D., Verheyen, K. (2000) *Streptococcus equi* with truncated M-proteins isolated from outwardly healthy horses. *Microbiology-Uk*, 146:1361-1369.
- Chanter, N, Ward, C. L., Talbot, N. C., Flanagan, J. A., Binns, M., Houghton, S. B., Smith, K. C., Mumford, J. A. (1999) Recombinant hyaluronate associated protein as a protective immunogen against *Streptococcus equi* and *Streptococcus zooepidemicus* challenge in mice. *Microbial Pathogenesis*, 27(3):133-143.
- Chen, P. S., Toribibara, T. Y. and Warner, H. (1956). Microdetermination of phosphorus. *Analytical Chemistry*, 28:1756-1758.
- Child, R., Wehrly, T. D., Rockx-Brouwer, D., Dorward, D. W., Celli, J. (2010) Acid phosphatases do not contribute to the pathogenesis of type A *Francisella tularensis*. *Infection and Immunity*, 78(1):59-67.
- Cossart, P., and Jonquieres, R. (2000) Sortase, a universal target for therapeutic agents against Gram-positive bacteria? *Proceedings of the National Academy of Sciences of the United States of America*, 97(10):5013-5015.
- Cottrell, R. C. (1981) Phospholipase A2 from bee venom. *Methods in enzymology*, 71 Pt C:698-702.
- Cregg, K. M., Wilding, E. I. and Black, M T (1996) Molecular cloning and expression of the *sysB* gene encoding an essential type I signal peptidase from *Staphylococcus aureus*. *Journal of Bacteriology*, 178(19):5712-5718.

Cunningham, M. W. (2000) Pathogenesis of group A streptococcal infections. *Clinical Microbiology Reviews*, 13(3):470.

Das, A., Asatryan, L., Reddy, M. A., Wass, C. A., Stins, M. F., Joshi, S., Bonventre, J. V., Kim, K. S. (2001) Differential role of cytosolic phospholipase A(2) in the invasion of brain microvascular endothelial cells by *Escherichia coli* and *Listeria monocytogenes*. *Journal of Infectious Diseases*, 184(6):732-737.

Davidson, A., Traub-Dargatz, J. L., Magnuson, R., Ashley, H., Irwin, V., Newton, R., Waller, A., Smith, K., Callan, R. J., Meehan, M., Owen, P., Salman, M. (2008) Lack of correlation between antibody titers to fibrinogen-binding protein of *Streptococcus equi* and persistent carriers of strangles. *Journal of Veterinary Diagnostic Investigation*, 20(4):457-462.

de Keyzer, J., van der Does, C. and Driessen, A. J. M. (2003) The bacterial translocase: a dynamic protein channel complex. *Cellular and Molecular Life Sciences*, 60(10):2034-2052.

Deangelis, P. L., Papaconstantinou, J. and Weigel, P. H. (1993) Molecular cloning, identification, and sequence of the hyaluronan synthase gene from group A *Streptococcus pyogenes*. *Journal of Biological Chemistry*, 268(26):19181-19184.

Dellabona, P., Peccoud, J., Kappler, J., Marrack, P., Benoist, C., Mathis, D. (1990) Superantigens interact with MHC class-II molecules outside of the antigen groove. *Cell*, 62(6):1115-1121.

Dennis, E. A. (2000) Phospholipase A(2) in eicosanoid generation. *American Journal of Respiratory and Critical Care Medicine*, 161(2):S32-S35.

- Dilks, K., Rose, R. W., Hartmann, E., Pohlschroder, M. (2003) Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *Journal of Bacteriology*, 185(4):1478-1483.
- Doery, H. M., Cheyne, I. M., Gulasekharan, J., Magnusson, B. J. (1963) A phospholipase in staphylococcal toxin which hydrolyses sphingomyelin. *Nature*, 198(488):1091.
- Dougherty, B. A., and Vanderijn, I. (1994) Molecular characterization of Hasa from an operon required for hyaluronic acid synthesis in group A *Streptococcus*. *Journal of Biological Chemistry*, 269(1):169-175.
- Dramsi, S., Magnet, S., Davison, S., Arthur, M. (2008) Covalent attachment of proteins to peptidoglycan. *Fems Microbiology Reviews*, 32(2):307-320.
- Dumon-Seignovert, L., Cariot, G. and Vuillard, L. (2004) The toxicity of recombinant proteins in *Escherichia coli*: A comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein expression and Purification*, 37(1):203-206.
- Du Plessis, D. J. F., Nouwen, N. and Driessen, A. J. M. (2011) The Sec translocase. *Biochimica Et Biophysica Acta-Biomembranes*, 1808(3):851-865.
- Du Plessis, E. M., Theron, J., Joubert, J., Joubert, L., Lotter, T., Watson, T. G. (2002) Characterization of a phosphatase secreted by *Staphylococcus aureus* strain 154, a new member of the bacterial class C family of nonspecific acid phosphatases. *Systematic and Applied Microbiology*, 25(1):21-30.
- Dvorak, H. F., Brockman, R. W. and Heppel, L. A. (1967). Purification and properties of two acid phosphatase fractions isolated from osmotic shock fluid of *Escherichia coli*. *Biochemistry*, 6(6):1743-51.

Edwards, A. T., Roulson, M. and Ironside, M. J. (1988) A milk-borne outbreak of serious infection due to *Streptococcus zooepidemicus* (Lancefield group C). *Epidemiology and Infection*, 101(1):43-51.

Egan, S. A., Kurian, D., Ward, P. N., Hunt, L., Leigh, J. A. (2010) Identification of Sortase A (SrtA) Substrates in *Streptococcus uberis*: Evidence for an Additional Hexapeptide (LPXXXD) Sorting Motif. *Journal of Proteome Research*, 9(2):1088-1095.

Emr, S. D., Hanleyway, S. and Silhavy, T. J. (1981) Suppressor mutations that restore export of a protein with a defective signal sequence. *Cell*, 23(1):79-88.

Evers, W. D. (1968) Effect of furaltadone on stranglers in horses. *Journal of the American Veterinary Medical Association*, 152(9):1394.

Facey, S. J., and Kuhn, A. (2010) Biogenesis of bacterial inner-membrane proteins. *Cellular and Molecular Life Sciences*, 67(14):2343-2362.

Facklam, R. (2002) What Happened to the Streptococci: Overview of Taxonomic and Nomenclature Changes. *Clinical. Microbiology. Review*, 15(4):613-630.

Fan, H. J., Wang, Y. S., Tang, F. Y., Lu, C. P. (2008) Determination of the mimic epitope of the M-like protein adhesin in swine *Streptococcus equi* subsp *zooepidemicus*. *Bmc Microbiology*, 8:9.

Felts, R. L., Reilly, T. J. and Tanner, J. J. (2005) Crystallization of AcpA, a respiratory burst-inhibiting acid phosphatase from *Francisella tularensis*. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 1752(1):107-110.

- Fernandez, E., Blume, V., Garrido, P., Collins, M. D., Mateos, A., Dominguez, L. (2005) *Streptococcus equi* subsp. *ruminatorum* subsp. nov., isolated from mastitis in small ruminants. *International Journal of Systematic and Evolutionary Microbiology*, 55:545-545.
- Fischetti, V. A., Pancholi, V. and Schneewind, O. (1990) Conservation of a hexapeptide sequence in the anchor region of surface proteins from Gram-positive cocci. *Molecular Microbiology*, 4(9):1603-1605.
- Flock, M., Jacobsson, K., Frykberg, L., Hirst, T. R., Franklin, A., Guss, B., Flock, J. (2004) Recombinant *Streptococcus equi* Proteins Protect Mice in Challenge Experiments and Induce Immune Response in Horses. *Infection and Immunity*, 72(6):3228-3236.
- Florindo, H. F., Pandit, S., Goncalves, L. M. M., Alpar, H. O., Almeida, A. J. (2010) Surface modified polymeric nanoparticles for immunisation against equine strangles. *International Journal of Pharmaceutics* In Press, Accepted Manuscript.
- Gaillard, J. L., Berche, P., Frehel, C., Gouin, E., Cossart, P. (1991) Entry of *L. monocytogenes* into cells mediated by internalin, a repeat protein reminiscent of surface-antigens from Gram-positive cocci. *Cell*, 65(7):1127-1141.
- Garcia, J. L., Sanchez-Beato, A. R., Medrano, F. J., Lpoez, R. (1998) Versatility of choline-binding domain. *Microbial Drug Resistance-Mechanisms Epidemiology and Disease*, 4(1):25-36.
- Gase, K., Liu, G. W., Bruckmann, A., Steiner, K., Ozegowski, J., Malke, H. (1997) The LppC gene of *Streptococcus equisimilis* encodes a lipoprotein that is homologous to the e(P4) outer membrane protein from *Haemophilus influenzae*. *Medical Microbiology and Immunology*, 186(1):63-73.

Ghannoum, M. A. (2000) Potential role of phospholipases in virulence and fungal pathogenesis. *Clinical Microbiology Reviews*, 13(1):122.

Ghomashchi, F., Lin, Y., Hixon, M. S., Yu, B. Z., Annand, R., Jain, M. K., Gelb, M. H. (1998) Interfacial recognition by bee venom phospholipase A(2): Insights into nonelectrostatic molecular determinants by charge reversal mutagenesis. *Biochemistry*, 37(19):6697-6710.

Gibbons, H. S., Wolschendorf, F., Abshire, M., Niederweis, M., Braunstein, M. (2007) Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. *Journal of Bacteriology*, 189(14):5090-5100.

Gibson, C. M. and Caparon, M.G. (1996). Insertional inactivation of *Streptococcus pyogenes* *sod* suggests that pftF is regulated in response to a superoxide signal. *Journal of Bacteriology*, 178 (15). 4688-95.

Giesbrecht, P., Wecke, J. and Reinicke, B. (1976) On the morphogenesis of cell wall of *Staphylococci*. *International Review of Cytology*, 44:225-318.

Golovan, S., Wang, G. R., Zhang, J., Forsberg, C. W. (2000) Characterization and overproduction of the *Escherichia coli* appA encoded bifunctional enzyme that exhibits both phytase and acid phosphatase activities. *Canadian Journal of Microbiology*, 46(1):59-71.

Gralnick, J. A., Vali, H., Lies, D. P., Newman, D. K. (2006) Extracellular respiration of dimethyl sulfoxide by *Shewanella oneidensis* strain MR-1. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4669-4674.

Green, B. A., Farley, J. E., Quinn, T., Deich, R. A., Zlotnick, G. W. (1991) The e(P4) outer membrane protein of *Haemophilus influenzae* - Biologic activity of anti-E serum and cloning and sequencing of the structural gene. *Infection and Immunity*, 59(9):3191-3198.

Guruprasad, K., Reddy, B. V. B. and Pandit, M. W. (1990) Correlation between stability of a protein and its dipeptide composition - A novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Engineering*, 4(2):155-161.

Guss, B., Flock, M., Frykberg, L., Waller, A. S., Robinson, C., Smith, K. C., Flock, J. I. (2009) Getting to Grips with Strangles: An Effective Multi-Component Recombinant Vaccine for the Protection of Horses from *Streptococcus equi* Infection. *Plos Pathogens*, 5(9).

Hamilton, A., Harrington, D. and Sutcliffe, I. C. (2000) Characterization of acid phosphatase activities in the equine pathogen *Streptococcus equi*. *Systematic and Applied Microbiology*, 23(3):325-329.

Hamilton, A., Robinson, C., Sutcliffe, I. C., Slater, J., Maskell, D. J., Davis-Poynter, Smith, K., Waller, A., Harrington, D. J. (2006) Mutation of the maturase lipoprotein attenuates the virulence of *Streptococcus equi* to a greater extent than does loss of general lipoprotein lipidation. *Infection and Immunity*, 74(12):6907-6919.

Harrington, D. J., Greated, J. S., Chanter, N., Sutcliffe, I. C. (2000) Identification of lipoprotein homologues of pneumococcal PsaA in the equine pathogens *Streptococcus equi* and *Streptococcus zooepidemicus*. *Infection and Immunity*, 68(10):6048-6051.

Harrington, D. J., Sutcliffe, I. C. and Chanter, N. (2002) The molecular basis of *Streptococcus equi* infection and disease. *Microbes and Infection*, 4(4):501-510.

Heather, Z., Holden, M. T. G., Steward, K. F., Parkhill, J., Song, L. J., Challis, G. L., Robinson, C., Davis-Poynter, N., Waller, A. S. (2008) A novel streptococcal integrative conjugative element involved in iron acquisition. *Molecular Microbiology*, 70(5):1274-1292.

Hienz, S. A., Schennings, T., Heimdahl, A., Flock, J. I. (1996) Collagen binding of *Staphylococcus aureus* is a virulence factor in experimental endocarditis. *Journal of Infectious Diseases*, 174(1):83-88.

Hoffman, A. M., Staempfli, H. R., Prescott, J. F., Viel, L. (1991). Field-evaluation of a commercial M-protein vaccine against *Streptococcus equi* infection in foals. *American Journal of Veterinary Research*, 52(4):589-592.

Hoffman, K. and Stoffel, W. (1993). TMbase - A database of membrane spanning proteins segments. *Journal of Biological Chemistry*, 374, 166.

Holden, M. T. G., Heather, Z., Paillot, R., Steward, K. F., Webb, K., Ainslie, F., Jourdan, T., Bason, N. C., Holroyd, N. E., Mungall, K., Quail, M. A., Sanders, M., Willey, D., Brooks, K., Aanensen, D. M., Spratt, B. G., Jolley, K. A., Maiden, M. C. J., Kehoe, M., Chanter, N., Bentley, S. D., Robinson, C., Maskell, D. J., Parkhill, J., Waller, A. S. (2009) Genomic Evidence for the Evolution of *Streptococcus equi*: Host Restriction, Increased Virulence, and Genetic Exchange with Human Pathogens. *Plos Pathogens*, 5(3).

Hong, K. (2005) Identification and characterization of a novel fibronectin-binding protein gene from *Streptococcus equi* subspecies *zooepidemicus* strain VTU211. *Fems Immunology and Medical Microbiology*, 45(2):231-237.

Hotomi, M., Ikeda, Y., Suzumoto, M., Yamauchi, K., Green, B. A., Zlotnick, G., Billal, D. S., Shimada, J., Fujihara, K., Yamanaka, N. (2005) A recombinant P4 protein of *Haemophilus influenzae* induces specific immune responses biologically active against

nasopharyngeal colonization in mice after intranasal immunization. *Vaccine*, 23(10):1294-1300.

Hulting, G., Flock, M., Frykberg, L., Lannergard, J., Flock, J., Guss, B. (2009) Two novel IgG endopeptidases of *Streptococcus equi*. *Fems Microbiology Letters*, 298(1):44-50.

Hutchings, M. I., Palmer, T., Harrington, D. J., Sutcliffe, I. C. (2009) Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em. *Trends in Microbiology*, 17(1):13-21.

Hyde, C. A. C. and Missailidis, S. (2009) Inhibition of arachidonic acid metabolism and its implication on cell proliferation and tumour-angiogenesis. *International Immunopharmacology*, 9(6):701-715.

Ijaz, M., Velineni, S. and Timoney, J. F. (2011) Selective pressure for allelic diversity in SeM of *Streptococcus equi* does not affect immunoreactive proteins SzPSe or Se18.9. *Infection, Genetics and Evolution*, 11(5):1159-1163.

Ikebe, T., Wada, A., Inagaki, Y., Sugama, K., Suzuki, R., Tanaka, D., Tamaru, A., Fukinaga, Y., Abe, Y., Shimizu, Y., Watanabe, H. (2002) Dissemination of the phage-associated novel superantigen gene speL in recent invasive and noninvasive *Streptococcus pyogenes* M3/T3 isolates in Japan. *Infection and Immunity*, 70(6):3227-3233.

Istivan, T. S. and Coloe, P. J. (2006) Phospholipase A in Gram-negative bacteria and its role in pathogenesis. *Microbiology-Sgm*, 152:1263-1274.

Ivens, P. A. S., Matthews, D., Webb, K., Newton, J. R., Steward, K., Waller, A. S., Robinson, C., Slater, J. D. (2011) Molecular characterisation of 'strangles' outbreaks in the UK: The use

of M-protein typing of *Streptococcus equi* ssp *equi*. Equine Veterinary Journal, 43(3):359-364.

Jackson, A. D., Rayner, C. F. J., Dewar, A., Cole, P. J., Wilson, R. (1996) A human respiratory tissue organ culture incorporating an air interface. American Journal of Respiratory and Critical Care Medicine, 153(3):1130-1135.

Jacobs, A. A. C., Goovaerts, D., Nuijten, P. J. M., Theelen, R. P. H., Hartford, O. M. (2000). Investigations towards an efficacious and safe strangles vaccine: submucosal vaccination with a live attenuated *Streptococcus equi*. Veterinary Record, 147(20):563-567.

Jadoun, J., Burstein, E., Hanski, E., Sela, S. (1997) Proteins M6 and F1 are required for efficient invasion of group A streptococci into cultured epithelial cells. Advances in Experimental Medicine and Biology, 418:511-515.

Jakubovics, N., Brittan, J. L., Dutton, L. C., Jenkinson, H. F. (2009). Multiple adhesin proteins on the cell surface of *Streptococcus gordonii* are involved in adhesion to human fibronectin. Microbiology-Sgm, 155:3572-3580.

Johnston, J. W., Myers, L. E., Ochs, M. A., Benjamin, W. H., Briles, D. E., Hollingshead, S. K. (2004) Lipoprotein PsaA in virulence of *Streptococcus pneumoniae*: Surface accessibility and role in protection from superoxide. Infection and Immunity, 72(10):5858-5867.

Jonquieres, R., Biernes, H., Fiedler, F., Gounon, P., Cossart, P. (1999) Interaction between the protein InIB of *Listeria monocytogenes* and lipoteichoic acid: a novel mechanism of protein association at the surface of Gram-positive bacteria. Molecular Microbiology, 34(5):902-914.

- Jorm, L. R. (1990) Strangles in horse studs - incidence, risk factors and effects of vaccination. *Australian Veterinary Journal*, 67(12):436-439.
- Jorm, L. R., Love, D. N., Bailey, G. D., McKay, G. M., Briscoe, D. A. (1994) Genetic structure of populations of beta-hemolytic lancefieldgroup C streptococci from horses and their association with disease. *Research in Veterinary Science*, 57(3):292-299.
- Jovel, S. R., Kumagai, T., Danshiitsoodol, N., Matoba, Y., Nishimura, M., Sugiyama, M. (2006) Purification and characterization of the second *Streptomyces* phospholipase A2 refolded from an inclusion body. *Protein Expression and Purification*, 50(1):82-88.
- Karlstrom, A., Jacobsson, K., Flock, M., Flock, J. I., Guss, B. (2004) Identification of a novel collagen-like protein, SclC, in *Streptococcus equi* using signal sequence phage display. *Veterinary Microbiology*, 104(3-4):179-188.
- Kelley, L. A and Sternberg, M. J. E. (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocols*, 4:363-371.
- Kelly, C., Bugg, M., Robinson, C., Mitchell, Z., Davis-Poynter, N., Newton, J. R., Jolley, K. A., Maiden, M. C. J., Waller, A. S. (2006) Sequence variation of the SeM gene of *Streptococcus equi* allows discrimination of the source of strangles outbreaks. *Journal of Clinical Microbiology*, 44(2):480-486.
- Kemp-Symonds, J., Kemble, T. and Waller, A. (2007) Modified live *Streptococcus equi* ('strangles') vaccination followed by clinically adverse reactions associated with bacterial replication. *Equine Veterinary Journal*, 39(3):284-286.
- Kitten, T., Munro, S. M., Macrina, F. L. (2000) Genetic characterization of a *Streptococcus mutans* LraI family operon and role in virulence. *Infection and Immunity*, 68(8):4441-4451.

Kocks, C., Gouin, E., Tabouret, M., Berche, P., Ohayon, H., Coassart, P. (1992) L-monoacylglycerol induced actin assembly requires the acta gene product, A surface protein. *Cell*, 68(3):521-531.

Lambeau, G., and M. Lazdunski (1999) Receptors for a growing family of secreted phospholipases A(2). *Trends in Pharmacological Sciences*, 20(4):162-170.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 15(5259):680-5.

Lancefield, R. C. (1962) Current knowledge of type-specific M antigens of group A streptococci. *Journal of immunology*, 89:307-13.

Lancefield, R. C. (1933) A serological differentiation of human and other groups of hemolytic streptococci. *Journal of Experimental Medicine*, 59:441-158.

Langton, S. R., and Cesareo, S. D.(1992) *Helicobacter pylori* associated phospholipase A2 activity - A factor in peptic ulcer production. *Journal of Clinical Pathology*, 45(3):221-224.

Lannergard, J., Flock, M., Johansson, S., Flock, J. I., Guss, B. (2005) Studies of fibronectin-binding proteins of *Streptococcus equi*. *Infection and Immunity*, 73(11):7243-7251.

Lannergard, J., Frykberg, L. and Guss, B. (2003) CNE, a collagen-binding protein of *Streptococcus equi*. *Fems Microbiology Letters*, 222(1):69-74.

Lannergard, J. and Guss, B. (2006) IdeE, an IgG-endopeptidase of *Streptococcus equi* ssp *equi*. *Fems Microbiology Letters*, 262(2):230-235.

Lansing, M., Lellig, S., Mausolf, A., Martini, I., Crescenzia, F., Oregan, M., Prehm, P. (1993) Hyaluronate synthase cloning and sequencing of the gene from *Streptococcus* sp. *Biochemical Journal*, 289:179-184.

Lapenta, D., Rubens, C., Chi, E., Cleary, P. P. (1994) Group A streptococci efficiently invade human respiratory epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America*, 91(25):12115-12119.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21):2947-8.

Laus, F., Preziuso, S., Spaterna, A., Beribe, F., Tesei, B., Cuteri, V. (2007) Clinical and epidemiological investigation of chronic upper respiratory diseases caused by beta-haemolytic Streptococci in horses. *Comparative Immunology Microbiology and Infectious Diseases*, 30(4):247-260.

Lee, J. Y. and Spicer, A. P. (2000) Hyaluronan: a multifunctional, megaDalton, stealth molecule. *Current Opinion in Cell Biology*, 12(5):581-586.

Lenz, L. L., Mohammadi, S., Geissler, A., Portnoy, D. A. (2003) SecA2-dependent secretion of autolytic enzymes promotes *Listeria monocytogenes* pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21):12432-12437.

Lindmark, H. and Guss, B. (1999) SFS, a novel fibronectin-binding protein from *Streptococcus equi*, inhibits the binding between fibronectin and collagen. *Infection and Immunity*, 67(5):2383-2388.

Lindmark, H., Jacobsson, K., Frykberg, L., Guss, B. (1996) Fibronectin-binding protein of *Streptococcus equi* subsp *zooepidemicus*. *Infection and Immunity*, 64(10):3993-3999.

Lindmark, H., Jonsson, P., Engvall, E. O., Guss, B. (1999) Pulsed-field gel electrophoresis and distribution of the genes *zag* and *fnz* in isolates of *Streptococcus equi*. Research in Veterinary Science, 66(2):93-99.

Lindmark, H., Nilsson, M. and Guss, B. (2001) Comparison of the fibronectin-binding protein FNE from *Streptococcus equi* subspecies *equi* with FNZ from *S. equi* subspecies *zooepidemicus* reveals a major and conserved difference. Infection and Immunity, 69(5):3159-3163.

Lineweaver, H. and Burk, D. (1934) The determination of enzyme dissociation constants. Journal of the American Chemical Society, 56(3):658-666.

Llewelyn, M. and Cohen, J. (2002) Superantigens: microbial agents that corrupt immunity. The Lancet Infectious Diseases, 2(3):156-162.

Lr, J. (1990) Strangles in horse studs: Incidence, risk factors and effect of vaccination. Australian Veterinary Journal, 67(12):436-439.

Maarek, Y. S., Jacovi, M., Shtalhaim, M., Ur, S., Zernik, D., Ben-Shaul, I. Z. (1998) Webcutter: a system for dynamic and tailorable site mapping. Computer Networks and ISDN Systems. 29 (8-13):1269-1279.

Macfarlane, R. G., Oakley, C. L., Anderson, C. G. (1941) Hemolysis and the production of opalescence in serum and lecithovitellin by the alpha toxin of *Clostridium perfringens*. Journal of pathological bacteriology, 52:99-103.

Madsen, J. J., Linderöth, L., Subramanian, A. K., Andresen, T. L., Peters, G. H. (2011) Secretory Phospholipase A(2) Activity toward Diverse Substrates. Journal of Physical Chemistry B, 115(21):6853-6861.

- Malke, H. (1998) Cytoplasmic membrane lipoprotein LppC of *Streptococcus equisimilis* functions as an acid phosphatase. *Applied and Environmental Microbiology*, 64(7):2439-2442.
- Mandlik, A., Swierczynski, A., Das, A., Ton-That, H. (2008) Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends in Microbiology*, 16(1):33-40.
- Marchandin, H., Jumas-Bilak, E., Boumzebra, A., Vidal, D., Jonquet, O., Corne, P. (2007) Fatal *Streptococcus equi* subsp *ruminantium* infection in a man. *Emerging Infectious Diseases*, 13(12):1964-1966.
- Marraffini, L. A., DeDent, L. A. and Schneewind, O. (2006) Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiology and Molecular Biology Reviews*, 70(1):192.
- Matoba, Y., Katsube, Y. and Sugiyama, M. (2002) The crystal structure of prokaryotic phospholipase A(2). *Journal of Biological Chemistry*, 277(22):20059-20069.
- Mazmanian, S, K., Liu, G., Ton-That, H., Schneewind, O. (1999) *Staphylococcus aureus* Sortase, an Enzyme that Anchors Surface Proteins to the Cell Wall. *Science*, 285(5428):760-763.
- Meehan, M., Burke, F. M., Macken, S., Owen, P. (2009) Localization of the equine IgG-binding domain in the fibrinogen-binding protein (FgBP) of *Streptococcus equi* subsp *equi*. *Microbiology-Sgm*, 155:2583-2592.

- Meehan, M., Nowlan, P. and Owen, P. (1998) Affinity purification and characterization of a fibrinogen-binding protein complex which protects mice against lethal challenge with *Streptococcus equi* subsp. *equi*. Microbiology, 144(4):993-1003.
- Meeusen, E. N. T., Walker, J., Peters, A., Pastoret, P. P., Jungersen, G. (2007) Current status of veterinary vaccines. Clinical Microbiology Reviews, 20(3):489-510.
- Mesnager, S., Fontaine, T., Mignot, T., Delepiere, M., Mock, M., Fouet, A. (2000) Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. Embo Journal, 19(17):4473-4484.
- Meyer, M. C., Rastogi, P., Beckett, C. S., McHowat, J. (2005). Phospholipase A(2) inhibitors as potential anti-inflammatory agents. Current Pharmaceutical Design. 11(10):1301-1312.
- Miroux, B. and Walker, J. E. (1996) Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. Journal of Molecular Biology, 260(3):289-298.
- Mitchell, A. M., and Mitchell, T. J. (2010) *Streptococcus pneumoniae*: virulence factors and variation. Clinical Microbiology and Infection, 16(5):411-418.
- Mohapatra, N. P., Balagopal, A., Soni, S., Schlesinger, L. S., Gunn, J. S. (2007) AcpA is a Francisella acid phosphatase that affects intramacrophage survival and virulence. Infection and Immunity, 75(1): 390-396
- Mohapatra, N. P., Soni, S., Rajaram, M. V. S., Dang, P., Reilly, T. J., El-Benna, J., Clay, C. D. Schlesinger, L. S., Gunn, J. S. (2010). Francisella Acid Phosphatases Inactivate the NADPH Oxidase in Human Phagocytes. Journal of Immunology 184(9):5141-5150

- Moore, B. O., Bryans, J. T. (1969) Antigenic classification of group C animal streptococci. *Journal of the American Veterinary Medical Association*, 155(2 PART 2):416-421.
- Novy, R. and Morris, B. (2001). Use of glucose to control basal expression in the pET system. *Innovations Biotechniques* 12:1-3.
- Muckle, C. A., and Gyles, C. L. (1983) Relation of lipid-content and exotoxin production to virulence of *Corynebacterium pseudotuberculosis*. *American Journal of Veterinary Research*, 44(6):1149-1153.
- Nagiec, M. J., Lei, B., Parker, S. K., Vasil, M. L., Matsumoto, M., Ireland, R. M., Beres, S. B., Hoe, N. P., Musser, J. M. (2004) Analysis of a novel prophage-encoded Group A *Streptococcus* extracellular phospholipase A(2). *Journal of Biological Chemistry*, 279(44):45909-45918.
- Navarre, W. W., Daefer, S. and Schneewind, O. (1996) Cell wall sorting of lipoproteins in *Staphylococcus aureus*. *Journal of Bacteriology*, 178(2):441-446.
- Navarre, W. W. Schneewind, O. (1994) Proteolytic cleavage and cell wall anchoring at the LPXTG motif of surface proteins in Gram positive bacteria. *Molecular Microbiology*, 14(1):115-121.
- Navarre, W. W. and Schneewind, O. (1999) Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiology and Molecular Biology Reviews* 63(1):174.

- Neville, B. A. and O'Toole, P. W. (2010) Probiotic Properties of *Lactobacillus salivarius*: Adhesive Properties of *L. salivarius* to Human IECs have been Characterized in vitro. *Future Microbiology*, 5(5):759-774.
- Newton, J. R., Wood, J. L. N., Dunn, K. A., DeBrauwere, M. N., Chanter, N. (1997) Naturally occurring persistent and asymptomatic infection of the guttural pouches of horses with *Streptococcus equi*. *Veterinary Record*, 140(4):84-90.
- Newton, R., Waller, A. and King, A. (2005) Investigation of suspected adverse reactions following strangles vaccination in horses. *Veterinary Record*, 156(9):291-292.
- Nickel, V., Prehm, S., Lansing, M., Mausolf, A., Podbielski, A., Deutscher, J., Prehm, P. (1998) An ectoprotein kinase of group C streptococci binds hyaluronan and regulates capsule formation. *Journal of Biological Chemistry*, 273(37):23668-23673.
- Nielsen, H., Engelbrecht, J., Brunak, S., VonHeijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering*, 10(1):1-6.
- Nishiyama, K., Mizushima, S. and Tokuda, H. (1993) A novel membrane-protein involved in protein translocation across the cytoplasmic membrane of *Escherichia coli*. *Embo Journal* 12(9):3409-3415.
- Nygaard, Tyler K., Liu, M., McClure, M. J., Lei, B. (2006) Identification and characterization of the heme-binding proteins SeShp and SeHtsA of *Streptococcus equi* subspecies *equi*. *BMC Microbiology*, 6:82
- Oki, M., Doi, O. and Nojima, S. (1972) Mutant of *Escherichia coli* K-12 deficient for detergent-resistant phospholipase A. *Journal of Bacteriology*, 110(3):864-9.

Oliver, D. B., and Beckwith, J. (1981) *Escherichia coli* mutant pleiotropically defective in the export of secreted proteins. *Cell*, 25(3):765-772.

Ou, Z. H., Felts, R. L., Reilly, T. J., Nix, J. C., Tanner, J. J. (2006) Crystallization of recombinant *Haemophilus influenzae* e (P4) acid phosphatase. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*, 62:464-466.

Ouyang, C., and Huang, T. F. (1984) Effect of the purified phospholipases A2 from snake and bee venoms on rabbit platelet function. *Toxicon*, 22(5):705-718.

Ozeri, V., Rosenshine, I., Mosher, D. F., Fassler, R., Hanski, E. (1998) Roles of integrins and fibronectin in the entry of *Streptococcus pyogenes* into cells via protein F1. *Molecular Microbiology*, 30(3):625-637.

Paetzel, M., Dalbey, R. E. and Strynadka, N. C. J. (1998) Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature*, 396(6707):186-190.

Paillot, R., Darby, A. C., Robinson, C., Wright, N. L., Steward, K. L., Anderson, E., Webb, K., Holden, M. T. G., Efstratiou, A., Broughton, K., Jolley, K. A., Priestnall, S. L., Campi, M. C. M., Hughes, M. A., Radford, A., Erles, K., Waller, A. S. (2010) Identification of Three Novel Superantigen-Encoding Genes in *Streptococcus equi* subsp *zooepidemicus*, *szeF*, *szeN*, and *szeP*. *Infection and Immunity*, 78(11):4817-4827.

Paillot, R., Robinson, C., Steward, K., Wright, N., Jourdan, T., Butcher, N., Heather, Z., Waller, A. S. (2010) Contribution of Each of Four Superantigens to *Streptococcus equi*-Induced Mitogenicity, Gamma Interferon Synthesis, and Immunity. *Infection and Immunity*, 78(4):1728-1739.

- Palacin, A., Parro, V., Geukens, N., Anne, J., Mellado, R. P. (2002) SipY is the *Streptomyces lividans* type I signal peptidase exerting a major effect on protein secretion. *Journal of Bacteriology*, 184(17):4875-4880.
- Pallen, M. J., Lam, A. C., Antonio, M., Dunbar, K. (2001) An embarrassment of sortases - a richness of substrates? *Trends in Microbiology*, 9(3):97-101.
- Parillo, F., Rossi, G., Busoni, V., Magi, G. E., Supplizi, A. V. (2009) Differentiation of glycans in equine guttural pouches. *Veterinary Journal*, 180(2):246-252.
- Park, S. H., Liu, G. P., Topping, T. B., Cover, W. H., Randall, L. L. (1988) Modulation of folding pathways of exposed proteins by the leader sequence. *Science*, 239(4843):1033-1035.
- Passariello, C., Schippa, S., Lori, P., Berlutti, F., Thaller, M. C., Rossolini, G. M. (2003) The molecular class C acid phosphatase of *Chryseobacterium meningosepticum* (OlpA) is a broad-spectrum nucleotidase with preferential activity on 5'-nucleotides. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 1648(1-2):203-209.
- Piche, C. A. (1984) Clinical observations on an outbreak of stranglers. *Canadian Veterinary Journal-Revue Veterinaire Canadienne*, 25(1):7-11.
- Poi, M. J., Tomaszewski, J. W., Yuan, C., Dunlap, C. A., Andersen, N. H., Gelb, M. H., Tsai, M. D. (2003) A Low-barrier Hydrogen Bond Between Histidine of Secreted Phospholipase A2 and a Transition State Analog Inhibitor. *Journal of Molecular Biology*, 329(5):997-1009.
- Poyart, C., Quesne, G., Coulon, S., Berche, P., Trieu-Cuot, P. (1998) Identification of streptococci to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase. *Journal of Clinical Microbiology*, 36(1):41-47.

Price, J. A. (2007) A colorimetric assay for measuring phospholipase A(2) degradation of phosphatidylcholine at physiological pH. *Journal of Biochemical and Biophysical Methods*, 70(3): 441-444.

Pusterla, N., Watson, J. L., Affolter, V. K., Magdesian, K. G., Wilson, W. D., Carlson, G. P. (2003) Purpura haemorrhagica in 53 horses. *Veterinary Record*, 153(4):118-121.

Rahman, O., Cummings, S. P., Harrington, D. J., Sutcliffe, I. C. (2008) Methods for the bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-positive bacteria. *World Journal of Microbiology & Biotechnology*, 24(11):2377-2382.

Rasko, D. A., Ravel, J., Okstad, O. A., Helgason, E., Cer, R. Z., Jiang, L. X., Shores, K. A., Fouts, D. E., Tourasse, N. J., Angiuoli, S. V., Kolonay, J., Nelson, W. C., Kolsto, A. B., Fraser, C. M., Read, T. D. (2004) The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1. *Nucleic Acids Research*, 32(3):977-988.

Reif, J.S., George, J. L. and Shideler, R. K. (1981) Recent developments in strangles research: observations on the carrier state and evaluation of a new vaccine. *Proceedings in American Association of Equine Practition*, 27:33-40.

Reilly, T. J., Baron, G. S., Nano, F. E., Kuhlenschmidt, M. S. (1996) Characterization and sequencing of a respiratory burst-inhibiting acid phosphatase from *Francisella tularensis*. *Journal of Biological Chemistry*, 271(18):10973-10983.

Reilly, T. J., Chance, D. D., Calcutt, M. J., Tanner, J. J., Felts, R. L., Waller, S. C., Henzi, M. T., Mawhinney, T. P., Ganjam, I. K., Fales, W. H. (2009) Characterization of a Unique Class C Acid Phosphatase from *Clostridium perfringens*. *Applied and Environmental Microbiology*, 75(11):3745-3754.

- Reilly, T. J., Chance, D. L. and Smith, A. L. (1999) Outer membrane lipoprotein e (P4) of *Haemophilus influenzae* is a novel phosphomonoesterase. *Journal of Bacteriology*, 181(21):6797-6805.
- Reilly, T. J., Felts, R. L., Henzi, M. T., Calcutt, M. J. (2006) Characterization of recombinant *Francisella tularensis* acid phosphatase A. *Protein Expression and Purification*, 45(1):132-141.
- Reilly, T. J., Green, B. A., Zlotnick, G. W., Smith, A. L. (2001) Contribution of the DDDD motif of H-influenzae e (P4) to phosphomonoesterase activity and heme transport. *FEBS Letters*, 494(1-2):19-23.
- Reilly, T. J., and Smith, A. L. (1999) Purification and characterization of a recombinant *Haemophilus influenzae* outer membrane phosphomonoesterase e (P4). *Protein Expression and Purification*, 17(3):401-409.
- Reilly, T. J., and Calcutt, M. J. (2004) The class C acid phosphatase of *Helicobacter pylori* is a 5' nucleotidase. *Protein Expression and Purification*, 33(1):48-56.
- Reynolds, L. J., Hughes, L. L. and Dennis, E. A. (1992). Analysis of human synovial fluid phospholipase A2 on short chain phosphatidylcholine-mixed micelles: Development of a spectrophotometric assay suitable for a microtiterplate reader
- Riggs, P. D., Derman, A. I. and Beckwith, J. (1988) A mutation affecting the regulation of a SecA-LacZ fusion defines a new Sec gene. *Genetics*, 118(4):571-579.
- Rosch, J. W., and Caparon, M. G. (2005) The ExPortal: an organelle dedicated to the biogenesis of secreted proteins in *Streptococcus pyogenes*. *Molecular Microbiology*, 58(4):959-968.

- Rossolini, G. M., Schippa, S., Riccio, M. L., Berlutti, F., Macaskie, L. E., Thaller, M. C. (1998) Bacterial nonspecific acid phosphohydrolases: physiology, evolution and use as tools in microbial biotechnology. *Cellular and Molecular Life Sciences*, 54(8):833-850.
- Takai, S., Yashiro, H., Ishii, C., Tsubaki, S., Anzai, T., Wada, R., Timoney, J. F. (2000) Detection of DNA restriction fragment polymorphisms in *Streptococcus equi*. *Journal of the British Veterinary Association*, (146) 159-161.
- Saha, A. K., Dowling, J. N., Lamarco, K. L., Das, S., Remaley, A. T., Olomu, N., Pope, M. T., Glew, R. H. (1985) Properties of an acid phosphatase from *Legionella micdadei* which blocks superoxide anion production by human neutrophils. *Archives of Biochemistry and Biophysics*, 243(1):150-160.
- Salton, M. R. J., Ghuyssen, J. M and Hakenbeck, R. (1994) Chapter 1 The bacterial cell envelope - a historical perspective. *In* New Comprehensive Biochemistry. Pp. 1-22: Elsevier.
- Scandell, C. J. and Kornberg, A. (1971) Membrane-bound phospholipase A₁ purified from *Escherichia coli*. *Biochemistry*, 10(24):4447-56.
- Schatz, P. J., and Beckwith, J. (1990) Genetic analysis of protein export in *Escherichia coli*. *Annual Review of Genetics*, 24:215-248.
- Schein, C. H. (1989) Production of Soluble Recombinant Proteins in Bacteria. *Nat Biotech*, 7(11):1141-1149.
- Schievella, A. R., Regier, M. K., Smith, W. L., Lin, L. L. (1995) Calcium-mediated translocation of cytosolic phospholipase A₂ to the nuclear envelope and endoplasmic reticulum. *Journal of Biological Chemistry*, 270(51):30749-30754.

Schleifer, K. (1973) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriology Review*, 37(2):258.

Schmiel, D. H., Wagar, E., Karamanou, L., Weeks, D., Miller, V. L. (1998) Phospholipase A of *Yersinia enterocolitica* contributes to pathogenesis in a mouse model. *Infection and Immunity*, 66(8):3941-3951.

Schunder, E., Adam, P., Higa, F., Remer, K. A., Lorenza, U., Bender, J., Schulz, T., Flieger, A., Steinert, M., Heuner, K. (2010) Phospholipase PlaB is a new virulence factor of *Legionella pneumophila*. *International Journal of Medical Microbiology*, 300(5):313-23.

Schwarz-Linek, U., Werner, J. M., Pickford, A. R., Gurusiddappa, S., Kim, J. H., Pika, E. S., Briggs, J. A. G., Gough, T. S., Hook, M., Campbell, I. D., Potts, J. R. (2003) Pathogenic bacteria attach to human fibronectin through a tandem beta-zipper. *Nature*, 423(6936):177-181.

Scott, J. R., and Barnett, T. C. (2006) Surface proteins of gram-positive bacteria and how they get there. *Annual Review of Microbiology*, 60:397-423.

Sekar, K., Yu, B. Z., Rodgers, J., Lutton, J., Liu, X. H., Chen, X., Tsai, M. D., Jain, M. K., Sundaralingam, M. (1997) Phospholipase A(2) engineering. Structural and functional roles of the highly conserved active site residue aspartate-99. *Biochemistry*, 36(11):3104-3114.

Shaw, M. K and Ingraham, J. L. (1967). Synthesis of macromolecules by *Escherichia Coli* near minimal temperature for growth. *Journal of bacteriology*. 94(1)157

Sheoran, A. S., Artiushin, S. and Timoney, J. F. (2002) Nasal mucosal immunogenicity for the horse of a SeM peptide of *Streptococcus equi* genetically coupled to cholera toxin. *Vaccine*, 20(11-12):1653-1659.

Sherman, J. M. (1937) The Streptococci. Bacteriological reviews, 1(1):3-97.

Shimizu, T., Ohtani, K., Hirakawa, H., Ohshima, K., Yamashita, A., Shiba, T., Ogasawara, N., Hattori, M., Kuhara, S., Hayashi, H. (2002) Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. Proceedings of the National Academy of Sciences of the United States of America, 99(2):996-1001.

Silhavy, T. J., Kahne, D. and Walker, S. (2010) The Bacterial Cell Envelope. Cold Spring Harbor Perspectives in Biology, 2(5).

Simpson, W. J., Musser, J. M. and Cleary, P. P. (1992) Evidence consistent with horizontal transfer of the gene (emm12) encoding serotype M12 protein between group A and group G pathogenic streptococci. Infection and Immunity, 60(5):1890-1893.

Sinha, B., Francois, P. P., Nusse, O., Foti, M., Hartford, O. M., vaudaux, P., Foster, T. J., Lew, D. P., Herrmann, M., Krause, K. H. (1999) Fibronectin-binding protein acts as *Staphylococcus aureus* invasin via fibronectin bridging to integrin $\alpha(5)\beta(1)$. Cellular Microbiology, 1(2):101-117.

Sitkiewicz, I., Nagiec, M. J., Sumby, P., Butler, S. D., Cywes-Bently, C., Musser, J. M. (2006) Emergence of a bacterial clone with enhanced virulence by acquisition of a phage encoding a secreted phospholipase A(2). Proceedings of the National Academy of Sciences of the United States of America, 103(43):16009-16014.

Sitkiewicz, I., K. E. Stockbauer, and J. M. Musser (2007) Secreted bacterial phospholipase A(2) enzymes: better living through phospholipolysis. Trends in Microbiology, 15(2):63-69.

Sjoquist, J., Movitz, J., Johansson, I. B., Hjelm, H. (1972) Localization of protein A in bacteria. European Journal of Biochemistry, 30(1):190.

Slesarev, A. I., Mezhevaya, K. V., Markarova, K. S., Polushin, N. N., Shcherbinina, O. V., Shakhova, V. V., Belova, G. I., Aravind, L., Natale, D. A., Rogozin, I. B., Tatusov, R. L., Wolf, Y. I., Stetter, K. O., Malykh, A. G., Koonin, E. V., Kozyavkin, S. A. (2002) The complete genome of hyperthermophile *Methanopyrus kandleri* AV19 and monophyly of archaeal methanogens. Proceedings of the National Academy of Sciences of the United States of America, 99(7):4644-4649.

Smeesters, P. R., McMillan, D. J. and Sriprakash, K. S. (2010) The streptococcal M protein: a highly versatile molecule. Trends in Microbiology, 18(6):275-282.

Smoot, L. M., McCormick, J. K., Smoot, J. C., Hoe, N. P., Strickland, I., Cole, R. L., Barbian, K. D., Earhart, C. A., Ohlendorf, D. H., Veasy, L. G., Hill, H. R., Leung, D. Y. M., Schlievert, P. M., Musser, J. M. (2002) Characterization of two novel pyrogenic toxin superantigens made by an acute rheumatic fever clone of *Streptococcus pyogenes* associated with multiple disease outbreaks. Infection and Immunity, 70(12):7095-7104.

Snowden, M. A., and Perkins, H. R. (1990) Peptidoglycan cross-linking in *staphylococcus aureus* - An apparent random polymerization process. European Journal of Biochemistry, 191(2):373-377.

Sriskandan, S., Faulkner, L. and Hopkins, P. (2007) *Streptococcus pyogenes*: Insight into the function of the streptococcal superantigens. International Journal of Biochemistry & Cell Biology, 39(1):12-19.

Stevens, D. L., Tweten, R. K., Awad, M. M., Rood, J. I., Bryant, A. E. (1997) Clostridial gas gangrene: Evidence that alpha and theta toxins differentially modulate the immune response and induce acute tissue necrosis. Journal of Infectious Diseases, 176(1):189-195.

Studier, F. W. (1991) Use of bacteriophage-T7 lysozyme to improve an inducible T7 expression system. *Journal of Molecular Biology*, 219(1):37-44.

Sugiyama, M., Ohtani, K., Izuhara, M., Koike, T., Suzuki, K., Imamura, S., Misaki, H. (2002) A novel prokaryotic phospholipase A(2) - Characterization, gene cloning, and solution structure. *Journal of Biological Chemistry*, 277(22):20051-20058.

Sutcliffe, I. C. (2010) A phylum level perspective on bacterial cell envelope architecture. *Trends in Microbiology*, 18(10):464-470.

Sweeney, C. R., Timoney, J. E., Newton, J. R., Hines, M. T. (2005) *Streptococcus equi* infections in horses: Guidelines for treatment, control, and prevention of strangles. *Journal of Veterinary Internal Medicine*, 19(1):123-134.

Takai, S., Anzai, T., Yashiro, C., Ishii, S., Tsubaki, R.(2000) Detection of DNA restriction fragment polymorphism in *Streptococcus equi*. *Veterinary Records* 146:159-161

Tannaes, T., Grav, H. T. and Bukholm, G. (2000) Lipid profiles of *Helicobacter pylori* colony variants. *Apmis* 108(5):349-356.

Thaller, M. C., Berlutti, F., Schippa, S., Lori, P., Passariello, C., Rossolini, G. M. (1995) Heterogeneous patterns of acid-phosphatases containing low-molecular-mass polypeptides in members of the family Enterobacteriaceae. *International Journal of Systematic Bacteriology* 45 (2): 255-261.

Thaller, M. C., Berlutti, F., Schippa, S., Lombardi, G., Rossolini, G. M. (1994) Characterization and sequence of PhoC, the principal phosphate-irrepressible acid-phosphatase of *Morganella morganii*. *Microbiology-UK* 140:1341-1350

- Thaller, M. C., Lombardi, G., Berlutti, F., Schippa, S., Rossolini, G. M. (1995) Cloning and characterization of the napa acid-phosphatase phosphotransferase of *Morganella morganii* - Identification of a new family of bacterial acid-phosphatase-encoding genes. Microbiology-UK 141:147-154
- Thaller, M. C., Schippa, S and Rossolini, G. M. (1998) Conserved sequence motifs among bacterial, eukaryotic, and archaeal phosphatases that define a new phosphohydrolase superfamily. Protein Science 7(7):1647-1652.
- Thompson, B, J., Widdick, D. A., Hicks, M. G., Chandra, G., Sutcliffe, I. C., Palmer, T., Hutchings, M. I. (2010) Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium *Streptomyces coelicolor*. Molecular Microbiology 77(4):943-957.
- Timoney, J. (1999) Equine Strangles. Proceedings of the annual convention of the AEEP 45:31-36.
- Timoney, J. F. (2004) The pathogenic equine streptococci. Veterinary Research, 35(4):397-409.
- Timoney, J. F., Artiushin, S. C. (1997) Detection of *Streptococcus equi* in equine nasal swabs and washes by DNA amplification. Veterinary Record, 141(17):446-447.
- Timoney, J. F., Artiushin, S. C. and Boschwitz, J. S. (1997) Comparison of the sequences and functions of *Streptococcus equi* M-like proteins SeM and SzPSe. Infection and Immunity, 65(9):3600-3605.
- Timoney, J. F., DeNegri, R., Sheoran, A., Forster, N. (2010) Affects of N-terminal variation in the SeM protein of *Streptococcus equi* on antibody and fibrinogen binding. Vaccine 28(6):1522-1527.

Timoney, J. F., Qin, A., Muthupalani, S., Artiushin, S. (2007) Vaccine potential of novel surface exposed and secreted proteins of *Streptococcus equi*. *Vaccine*, 25(30):5583-5590.

Timoney, J. F., Walker, J., Zhou, M., Ding, J. B.(1995) Cloning and sequence analysis of a protective M-like protein gene from *Streptococcus equi* subsp *zooepidemicus*. *Infection and Immunity* 63(4):1440-1445.

Timoney, J. F., Yang, J., Liu, J., Merant, C. (2008) IdeE reduces the bactericidal activity of equine neutrophils for *Streptococcus equi*. *Veterinary Immunology and Immunopathology* 122(1-2):76-82.

Tiwari, R., Qin, A. P., Artiushin, S., Timoney, J. F. (2007) Se18.9, an anti-phagocytic factor H binding protein of *Streptococcus equi*. *Veterinary Microbiology*, 121(1-2):105-115.

Tjalsma, H., Noback, M. A., Bron, S., Venema, G., Yamane, K., VanDijl. J. M. (1997) *Bacillus subtilis* contains four closely related type I signal peptidases with overlapping substrate specificities - Constitutive and temporally controlled expression of different sip genes. *Journal of Biological Chemistry*, 272(41):25983-25992.

Tlapak-Simmons, V. L., Baggenstoss, B. A., Clyne, T., Weigel, P. H. (1999) Purification and lipid dependence of the recombinant hyaluronan syntheses from *Streptococcus pyogenes* and *Streptococcus equisimilis*. *Journal of Biological Chemistry*, 274(7):4239-4245.

Turner, C. E., Jurupati, P., Wiles, S., Edwards, R. J., Sriskandan, S. (2009) Impact of immunization against SpyCEP during invasive disease with two streptococcal species: *Streptococcus pyogenes* and *Streptococcus equi*. *Vaccine*, 27(36):4923-4929.

Tuteja, R. (2005) Type I signal peptidase: An overview. *Archives of Biochemistry and Biophysics*, 441(2):107-111.

- Uerkvitz, W (1988) Periplasmic nonspecific acid phosphatase II from *salmonella typhimurium* LT2 - Crystallization, detergent, reactivation and phosphotransferase activity. *Journal of Biological Chemistry*, 263(30):15823-15830.
- Van Roosmalen, M. L., Geukens, N., Jongbloed, J. D. H., Tjalsma, H., Dubois, J. Y. F., Bron, S., Van Dijk, J. M., Anne, J. (2004) Type I signal peptidases of gram-positive bacteria. *Biochimica Et Biophysica Acta-Molecular Cell Research*, 1694(1-3):279-297.
- Verheyen, K., Newton, J. R., Talbot, N. C., de Brauwere, M. N., Chanter, N. (2000) Elimination of guttural pouch infection and inflammation in asymptomatic carriers of *Streptococcus equi*. *Equine Veterinary Journal*, 32(6):527-532.
- Von Pawel-Rammingen, U., Johansson, B. P. and Bjorck, L. (2002) IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. *Embo Journal*, 21(7):1607-1615.
- Vonheijne, G. and Abrahmsen, L. (1989) Species-specific variation in signal peptide design-implication for protein secretion in foreign hosts. *FEBS Letters*, 244(2):439-446.
- Walker, D. H., Feng, H. M. and Popov, V. L. (2001) Rickettsial phospholipase A(2) as a pathogenic mechanism in a model of cell injury by typhus and spotted fever group rickettsiae. *American Journal of Tropical Medicine and Hygiene*, 65(6):936-942.
- Waller, A., Flock, M., Smith, K., Robinson, C., Mitchell, Z., Karlstrom, A., Lannergard, J., Bergman, R., Guss, B., Flock, J. I. (2007) Vaccination of horses against strangles using recombinant antigens from *Streptococcus equi*. *Vaccine*, 25(18):3629-3635.
- Waller, A. S., and Jolley, K. A. (2007) Getting a grip on strangles: Recent progress towards improved diagnostics and vaccines. *Veterinary Journal*, 173(3):492-501.

- Webb, K., Jolley, K. A., Mitchell, Z., Robinson, C., Newton, J. R., Maiden, M. C. J., Waller, A. (2008) Development of an unambiguous and discriminatory multilocus sequence typing scheme for the *Streptococcus zooepidemicus* group. *Microbiology*, 154:3016-3024.
- Whatmore, A. M. (2001) *Streptococcus pyogenes* sclB encodes a putative hypervariable surface protein with a collagen-like repetitive structure. *Microbiology*, 147:419-429.
- Wibawan, I. W. T., Pasaribu, F. H., Utama, I. H., Abdulmawjood, A., Lammler, C. (1999) The role of hyaluronic acid capsular material of *Streptococcus equi* subsp *zooepidemicus* in mediating adherence to HeLa cells and in resisting phagocytosis. *Research in Veterinary Science*, 67(2):131-135.
- Widdick, D. A., Hicks, M. G., Thimpson, B. J., Tschumi, A., Chandra, G., Sutcliffe, I. C., Bruelle, J. K., Sander, P., Palmer, T., Hutchings, M. I. (2011) Dissecting the complete lipoprotein biogenesis pathway in *Streptomyces scabies*. *Molecular Microbiology*, 80(5):1395-1412.
- Wood, W. B. (1966) Host specificity of DNA produced by *Escherichia coli*: Bacterial mutations affecting the restriction and modification of DNA. *Journal of Molecular Biology*, 16(1):118-133.
- Wooldridge, K. G. and Williams, P. H. (1993) Iron uptake mechanisms of pathogenic bacteria. *Fems Microbiology Reviews*, 12(4):325-348.
- Zhu, L. K., Maruvada, R., Sapirstein, A., Malik, K. U., Peters-Golden, M., Kim, K. S. (2010) Arachidonic Acid Metabolism Regulates *Escherichia coli* Penetration of the Blood-Brain Barrier. *Infection and Immunity*, 78(10):4302-4310.

Appendix I – Media, solutions and enzyme buffers

Media

Luria-Bertani (Liquid media, per 100 ml)

Tryptone	1.0g
NaCl ₂	1.0g
Yeast extract	0.5g
pH 7.0	

Luria-Bertani (Solid media, per 100 ml)

Tryptone	1.0g
NaCl ₂	1.0g
Yeast extract	0.5g
Agar (bacteriological agar N°1)	2.0g
pH 7.0	

Todd-Hewitt (Liquid media, per 100 ml)

Todd-Hewitt	3.64g
-------------	-------

pH 7.0

Todd-Hewitt (Solid Media, per 100 ml)

Todd-Hewitt	3.64g
-------------	-------

Agar (bacteriological agar N°1)	2.0g
---------------------------------	------

pH 7.0

Solutions

Coomassie blue protein stain

Brilliant Blue R250	1.0g
---------------------	------

Methanol	450ml
----------	-------

ddH ₂ O	450ml
--------------------	-------

Glacial acetic acid	100ml
---------------------	-------

Elution Buffer (per litre)

Na ₂ HPO ₄ (20mM)	2.84g
---	-------

NaCl ₂ (500 mM)	29.22g
----------------------------	--------

Imidazole (10 mM)	34.0g
-------------------	-------

pH 7.0-7.5

Equilibration buffer (1 litre)

Sodium acetate (100mM, pH 5.0) 8.203g

PBS- TWEEN (per litre)

KH₂PO₄ (1.4mM) 0.19g

Na₂HPO₄ (8.0mM) 1.136g

NaCl (140mM) 8.2g

KCl (2.7mM, pH 7.3) 0.202g

Tween 80 0.5ml

Phosphate colour reagent (50 ml)

Sulphuric acid (3M) 10ml

Ammonium molybdate (2.5%) 10ml

18.2MΩ H₂O 30ml

Ascorbic acid 1.0g

Renaturation buffer (1 litre)

Tris HCl (100mM, pH 7.0)	12.114g
MgSO ₄ (2mM)	0.24072g
ZnCl ₂ (0.05mM)	6.82mg
Triton X-100 (2%, v/v)	20ml

SDS-PAGE Buffer B (per 200ml)

Tris HCl (2M, pH 8.8)	150ml
SDS (10%)	8ml
18.2MΩ	42ml

SDS PAGE Destain (per litre)

Methanol	100ml
Glacial Acetic Acid	100ml
dH ₂ O	800ml

SDS Loading Buffer (per 10 ml)

Tris HCl (1M, pH 6.8)	0.6ml
-----------------------	-------

Glycerol (50%)	5.0ml
SDS (10%)	2.0ml
β Mercaptoethanol	0.6ml
Bromophenol Blue (1%)	0.1g
dH ₂ O	0.9ml

SDS-PAGE running buffer (1 litre)

Tris-HCl (pH 8.0)	30.3g
Glycine	144g
SDS	10g
pH 8.2	

Start Buffer (per litre)

Na ₂ HPO ₄ (20mM)	2.84g
NaCl ₂ (500 mM)	29.22g
Imadizole (10 mM)	0.68g
pH	7.0-7.5

STETS Buffer (per litre)

Sucrose	80.0g
Tris-HCl	6.06g
EDTA (0.5M)	100.0ml
Triton X-100	50.0ml
pH	8.0

sPLA₂ Assay buffer (per litre)

CaCl ₂ (10mM)	1.1098g
KCl (100mM)	7.45513g
Tris HCl (25mM, pH 7.5)	3.02185g
Triton X-100 (0.3mM)	176 µl
pH 7.5	

TAE Running buffer (50 x stock)(per litre)

Tris Base	242.0g
Glacial acetic acid (17.51M)	57.1ml
EDTA (0.5M, pH 8.0)	100.0ml

TE buffer (per litre)

Tris-Base (0.5M, pH 7.5)	20ml
--------------------------	------

EDTA (0.5M, pH 8.0)	20ml
---------------------	------

pH 7.5

Transfer Buffer (per litre)

Glycine	2.93g
---------	-------

Tris-HCl	5.82g
----------	-------

Methanol	200ml
----------	-------

ddH ₂ O	800ml
--------------------	-------

Western blot transfer buffer (per litre)

Tris base	5.82g
-----------	-------

Glycine	2.93g
---------	-------

Methanol	200ml
----------	-------

Xylene cyanol (6x) loading buffer (per 10ml)

Xylene cyanol	0.025g
---------------	--------

Glycerol

3.000g

Enzyme buffers

<u>Enzyme</u>	<u>Buffer Composition</u>
KOD (DNA polymerase)	Patented
NdeI	Buffer 4 (20 mM Tris-acetate, 10 mM Magnesium acetate, 50 mM Potassium acetate, 1 mM DTT, pH 7.9)
T4 DNA ligase	50 mM Tris-HCl pH 7.5, 10 mM MgCl ₂ , 10 mM DTT, 1mM ATP, 25 mg ml ⁻¹ acetylated BSA
XhoI	Buffer 4 (20 mM Tris-acetate, 10 mM Magnesium acetate, 50 mM Potassium acetate, 1 mM DTT, pH 7.9)
EcoRI	50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl ₂ , 1 mM Dithioehreitol, pH 7.9, 25°C

Appendix II – Commercial kits and equipment

Commercial kits

Qiagen:

QIAprep[®] Spin Miniprep Kit

QIAquick[®] Gel Extraction Kit

Qiagen[®] Plasmid Maxi Kit

<u>Qiagen[®] kit buffer</u>	<u>Buffer functions / components</u>
P1	Resuspension. 50 mM Tris-HCl pH 8, 10 mM EDTA, 100 µg/ml RNase A
P2	Lysis. 200 mM NaOH, 1% SDS (w/v)
P3	Neutralization. 3 M Potassium Acetate pH 5
N3	Neutralization. 25 – 50 % Guanidinium chloride, 10-25% acetic acid.
EB	Elution. Full components not available
PE	Wash. (Full components not available)
QG	Agarose gel dissolver.

	50-100% Guanidinium chloride, (full components not available)
PB	Wash. 25-50% Guanidinium chloride, 25-50% isopropanol. (Full components not available)
QBT	Equilibration. 750 mM NaCl, 50 mM MOPS pH 7, 15% isopropanol (v/v), 0.15 % Triton X-100 (v/v)
QC	Wash. 1 M NaCl, 50 mM Tris-HCl, pH 8.5, 15 % isopropanol (v/v)
QF	Elution. 1.25 M NaCl, 50 mM Tris-HCl, pH 8.5, 15 % isopropanol (v/v)

Caymen Chemical Company: sPLA₂ assay kit

<u>Caymen chemical company kit</u>	<u>Buffer functions / components</u>
sPLA ₂ Buffer	Buffer 25 mM Tris-HCl, pH7.5, containing 10 mM CaCl ₂ , 100 mM KCl, 0.3 mM Triton X-100
DTNB	Chromogen 10 mM DTNB, 0.4 M Tris-HCl
Substrate	DT-PC (1.66 mM) 25 mM Tris-HCl, pH7.5, containing 10 mM CaCl ₂ , 100 mM KCl, 0.3 mM Triton X-100

Control	Bee venom (100 µg/ml) 25 mM Tris-HCl, pH7.5, containing 10 mM CaCl ₂ , 100 mM KCl, 0.3 mM Triton X-100
---------	--

Equipment

Autoclaving

Sterilisation was achieved using a benchtop Prestige[®] Medical 2100 Classic autoclave.

Incubators

For the growth of bacteria either a Gallenkamp orbital or static incubator was used.

pH meter

When solutions and media were required at a set pH all adjustments were performed using a Jenway Ion Meter 3340 with calibration buffers at pH 4.0, 7.0 and 9.22.

Centrifugation

Centrifugation of volumes was performed using a small Sigma 1-15 bench top microcentrifuge.

SDS-PAGE gel kits

Electrophoresis of proteins was performed using a Bio-Rad Mini-Protean 3 Cell kit with power supplied via an E-C 570-90 E-C Apparatus Corporation power pack

Gel Documentation

Visualisation of agarose and SDS-PAGE gels was achieved using a Bio-Rad Gel Doc 2000 system and Quantity One™ software. Hard copies of the gel picture were printed via a Mitsubishi Video Copy processor (Model P91), with Mitsubishi thermal paper.

Sonication

Cell lysis was performed using a MSE Soniprep 150 ultra-sonication machine

Purification

For purification of protein an AKTA Prime plus FPLC system was used and was controlled by PrimeView Software.

Spectrophotometer

For all spectrophotometric measurements a UV Visible Helios α Spectronic Unicam spectrophotometer was used.

Microtitre plate reader

Measurements were made using a FL500 FAK microtitre plate reader.

Appendix III Crystal tray screens

PEG/Ion Screen

PEG/Ion Screen™

HR2-126 Reagent Formulation

Tube #	Salt	Tube #	Polymer	Tube #	pH ◊	
1.	0.2 M Sodium fluoride	1.	20% w/v Polyethylene glycol 3,350	1.	7.3	F^- Fluoride
2.	0.2 M Potassium fluoride	2.	20% w/v Polyethylene glycol 3,350	2.	7.3	Cl^- Chloride
3.	0.2 M Ammonium fluoride	3.	20% w/v Polyethylene glycol 3,350	3.	6.2	I^- Iodide
4.	0.2 M Lithium chloride	4.	20% w/v Polyethylene glycol 3,350	4.	6.8	
5.	0.2 M Magnesium chloride hexahydrate	5.	20% w/v Polyethylene glycol 3,350	5.	5.9	
6.	0.2 M Sodium chloride	6.	20% w/v Polyethylene glycol 3,350	6.	6.9	$\text{O}=\text{N}-\text{O}^-$ Nitrate
7.	0.2 M Calcium chloride dihydrate	7.	20% w/v Polyethylene glycol 3,350	7.	5.1	
8.	0.2 M Potassium chloride	8.	20% w/v Polyethylene glycol 3,350	8.	7.0	
9.	0.2 M Ammonium chloride	9.	20% w/v Polyethylene glycol 3,350	9.	6.3	
10.	0.2 M Sodium iodide	10.	20% w/v Polyethylene glycol 3,350	10.	7.0	$\text{O}=\text{S}-\text{C}\equiv\text{N}$ Thiocyanate
11.	0.2 M Potassium iodide	11.	20% w/v Polyethylene glycol 3,350	11.	7.0	
12.	0.2 M Ammonium iodide	12.	20% w/v Polyethylene glycol 3,350	12.	6.2	
13.	0.2 M Sodium thiocyanate	13.	20% w/v Polyethylene glycol 3,350	13.	6.9	
14.	0.2 M Potassium thiocyanate	14.	20% w/v Polyethylene glycol 3,350	14.	7.0	
15.	0.2 M Lithium nitrate	15.	20% w/v Polyethylene glycol 3,350	15.	7.1	
16.	0.2 M Magnesium nitrate hexahydrate	16.	20% w/v Polyethylene glycol 3,350	16.	5.9	
17.	0.2 M Sodium nitrate	17.	20% w/v Polyethylene glycol 3,350	17.	6.8	
18.	0.2 M Potassium nitrate	18.	20% w/v Polyethylene glycol 3,350	18.	6.8	
19.	0.2 M Ammonium nitrate	19.	20% w/v Polyethylene glycol 3,350	19.	6.2	
20.	0.2 M Magnesium formate dihydrate	20.	20% w/v Polyethylene glycol 3,350	20.	7.0	
21.	0.2 M Sodium formate	21.	20% w/v Polyethylene glycol 3,350	21.	7.2	
22.	0.2 M Potassium formate	22.	20% w/v Polyethylene glycol 3,350	22.	7.3	
23.	0.2 M Ammonium formate	23.	20% w/v Polyethylene glycol 3,350	23.	6.6	
24.	0.2 M Lithium acetate dihydrate	24.	20% w/v Polyethylene glycol 3,350	24.	7.9	
25.	0.2 M Magnesium acetate tetrahydrate	25.	20% w/v Polyethylene glycol 3,350	25.	7.9	
26.	0.2 M Zinc acetate dihydrate	26.	20% w/v Polyethylene glycol 3,350	26.	6.4	
27.	0.2 M Sodium acetate trihydrate	27.	20% w/v Polyethylene glycol 3,350	27.	8.0	
28.	0.2 M Calcium acetate hydrate	28.	20% w/v Polyethylene glycol 3,350	28.	7.5	
29.	0.2 M Potassium acetate	29.	20% w/v Polyethylene glycol 3,350	29.	8.1	
30.	0.2 M Ammonium acetate	30.	20% w/v Polyethylene glycol 3,350	30.	7.1	
31.	0.2 M Lithium sulfate monohydrate	31.	20% w/v Polyethylene glycol 3,350	31.	6.0	
32.	0.2 M Magnesium sulfate heptahydrate	32.	20% w/v Polyethylene glycol 3,350	32.	6.0	
33.	0.2 M Sodium sulfate decahydrate	33.	20% w/v Polyethylene glycol 3,350	33.	6.7	
34.	0.2 M Potassium sulfate	34.	20% w/v Polyethylene glycol 3,350	34.	6.8	
35.	0.2 M Ammonium sulfate	35.	20% w/v Polyethylene glycol 3,350	35.	6.0	
36.	0.2 M Sodium tartrate dibasic dihydrate	36.	20% w/v Polyethylene glycol 3,350	36.	7.3	
37.	0.2 M Potassium sodium tartrate tetrahydrate	37.	20% w/v Polyethylene glycol 3,350	37.	7.4	
38.	0.2 M Ammonium tartrate dibasic	38.	20% w/v Polyethylene glycol 3,350	38.	6.6	
39.	0.2 M Sodium phosphate monobasic monohydrate	39.	20% w/v Polyethylene glycol 3,350	39.	4.7	
40.	0.2 M Sodium phosphate dibasic dihydrate	40.	20% w/v Polyethylene glycol 3,350	40.	9.1	
41.	0.2 M Potassium phosphate monobasic	41.	20% w/v Polyethylene glycol 3,350	41.	4.8	
42.	0.2 M Potassium phosphate dibasic	42.	20% w/v Polyethylene glycol 3,350	42.	9.2	
43.	0.2 M Ammonium phosphate monobasic	43.	20% w/v Polyethylene glycol 3,350	43.	4.6	
44.	0.2 M Ammonium phosphate dibasic	44.	20% w/v Polyethylene glycol 3,350	44.	8.0	
45.	0.2 M Lithium citrate tribasic tetrahydrate	45.	20% w/v Polyethylene glycol 3,350	45.	8.4	
46.	0.2 M Sodium citrate tribasic dihydrate	46.	20% w/v Polyethylene glycol 3,350	46.	8.3	
47.	0.2 M Potassium citrate tribasic monohydrate	47.	20% w/v Polyethylene glycol 3,350	47.	8.3	
48.	0.2 M Ammonium citrate dibasic	48.	20% w/v Polyethylene glycol 3,350	48.	5.1	

◊ Measured pH at 25 °C

PEG Anion

Well number	Salt	Buffer	PEG
1	0.2 M Sodium fluoride		20% (w/v) PEG 3350
2	0.2 M Sodium bromide		20% (w/v) PEG 3350
3	0.2 M Sodium iodide		20% (w/v) PEG 3350
4	0.2 M Potassium thiocyanate		20% (w/v) PEG 3350
5	0.2 M Sodium nitrate		20% (w/v) PEG 3350
6	0.2 M Sodium formate		20% (w/v) PEG 3350
7	0.2 M Sodium acetate		20% (w/v) PEG 3350
8	0.2 M Sodium sulphate		20% (w/v) PEG 3350
9	0.2 M Sodium potassium tartrate		20% (w/v) PEG 3350
10	0.2 M Sodium potassium phosphate		20% (w/v) PEG 3350
11	0.2 M Sodium citrate		20% (w/v) PEG 3350
12	0.2 M Sodium malonate		20% (w/v) PEG 3350
13	0.2 M Sodium fluoride	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
14	0.2 M Sodium bromide	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
15	0.2 M Sodium iodide	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
16	0.2 M Potassium thiocyanate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
17	0.2 M Sodium nitrate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
18	0.2 M Sodium formate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
19	0.2 M Sodium acetate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
20	0.2 M Sodium sulphate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
21	0.2 M Sodium potassium tartrate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
22	0.2 M Sodium potassium phosphate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
23	0.2 M Sodium citrate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
24	0.2 M Sodium malonate	0.1 M Bis-tris propane pH 6.5	20% (w/v) PEG 3350
25	0.2 M Sodium fluoride	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
26	0.2 M Sodium bromide	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
27	0.2 M Sodium iodide	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350

28	0.2 M Potassium thiocyanate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
29	0.2 M Sodium nitrate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
30	0.2 M Sodium formate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
31	0.2 M Sodium acetate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
32	0.2 M Sodium sulphate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
33	0.2 M Sodium potassium tartrate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
34	0.2 M Sodium potassium phosphate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
35	0.2 M Sodium citrate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
36	0.2 M Sodium malonate	0.1 M Bis-tris propane pH 7.5	20% (w/v) PEG 3350
37	0.2 M Sodium fluoride	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
38	0.2 M Sodium bromide	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
39	0.2 M Sodium iodide	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
40	0.2 M Potassium thiocyanate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
41	0.2 M Sodium nitrate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
42	0.2 M Sodium formate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
43	0.2 M Sodium acetate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
44	0.2 M Sodium sulphate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
45	0.2 M Sodium potassium tartrate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
46	0.2 M Sodium potassium phosphate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
47	0.2 M Sodium citrate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350
48	0.2 M Sodium malonate	0.1 M Bis-tris propane pH 8.5	20% (w/v) PEG 3350

PEG/Cation

Well number		Buffer	PEG
1	0.2 M Sodium chloride	Acetate pH 5	20% (w/v) PEG 6000
2	0.2 M Ammonium chloride	Acetate pH 5	20% (w/v) PEG 6000
3	0.2 M Lithium chloride	Acetate pH 5	20% (w/v) PEG 6000
4	0.2 M Magnesium chloride	Acetate pH 5	20% (w/v) PEG 6000
5	0.2 M Calcium chloride	Acetate pH 5	20% (w/v) PEG 6000
6	0.01 M Zinc chloride	Acetate pH 5	20% (w/v) PEG 6000
7	0.2 M Sodium chloride	MES pH 6	20% (w/v) PEG 6000
8	0.2 M Ammonium chloride	MES pH 6	20% (w/v) PEG 6000
9	0.2 M Lithium chloride	MES pH 6	20% (w/v) PEG 6000
10	0.2 M Magnesium chloride	MES pH 6	20% (w/v) PEG 6000
11	0.2 M Calcium chloride	MES pH 6	20% (w/v) PEG 6000
12	0.01 M Zinc chloride	MES pH 6	20% (w/v) PEG 6000
13	0.2 M Sodium chloride	HEPES pH 7	20% (w/v) PEG 6000
14	0.2 M Ammonium chloride	HEPES pH 7	20% (w/v) PEG 6000
15	0.2 M Lithium chloride	HEPES pH 7	20% (w/v) PEG 6000
16	0.2 M Magnesium chloride	HEPES pH 7	20% (w/v) PEG 6000
17	0.2 M Calcium chloride	HEPES pH 7	20% (w/v) PEG 6000
18	0.01 M Zinc chloride	HEPES pH 7	20% (w/v) PEG 6000
19	0.2 M Sodium chloride	Tris pH 8	20% (w/v) PEG 6000
20	0.2 M Ammonium chloride	Tris pH 8	20% (w/v) PEG 6000
21	0.2 M Lithium chloride	Tris pH 8	20% (w/v) PEG 6000
22	0.2 M Magnesium chloride	Tris pH 8	20% (w/v) PEG 6000
23	0.2 M Calcium chloride	Tris pH 8	20% (w/v) PEG 6000
24	0.01 M Zinc chloride	Tris pH 8	20% (w/v) PEG 6000

Index

Index™

HR2-144 Reagent Formulation

Tube #	Salt	Tube #	Buffer ◊	Tube #	Precipitant
1.	None	1.	0.1 M Citric acid pH 3.5	1.	2.0 M Ammonium sulfate
2.	None	2.	0.1 M Sodium acetate trihydrate pH 4.5	2.	2.0 M Ammonium sulfate
3.	None	3.	0.1 M BIS-TRIS pH 5.5	3.	2.0 M Ammonium sulfate
4.	None	4.	0.1 M BIS-TRIS pH 6.5	4.	2.0 M Ammonium sulfate
5.	None	5.	0.1 M HEPES pH 7.5	5.	2.0 M Ammonium sulfate
6.	None	6.	0.1 M Tris pH 8.5	6.	2.0 M Ammonium sulfate
7.	None	7.	0.1 M Citric acid pH 3.5	7.	3.0 M Sodium chloride
8.	None	8.	0.1 M Sodium acetate trihydrate pH 4.5	8.	3.0 M Sodium chloride
9.	None	9.	0.1 M BIS-TRIS pH 5.5	9.	3.0 M Sodium chloride
10.	None	10.	0.1 M BIS-TRIS pH 6.5	10.	3.0 M Sodium chloride
11.	None	11.	0.1 M HEPES pH 7.5	11.	3.0 M Sodium chloride
12.	None	12.	0.1 M Tris pH 8.5	12.	3.0 M Sodium chloride
13.	None	13.	0.1 M BIS-TRIS pH 5.5	13.	0.3 M Magnesium formate dihydrate
14.	None	14.	0.1 M BIS-TRIS pH 6.5	14.	0.5 M Magnesium formate dihydrate
15.	None	15.	0.1 M HEPES pH 7.5	15.	0.5 M Magnesium formate dihydrate
16.	None	16.	0.1 M Tris pH 8.5	16.	0.3 M Magnesium formate dihydrate
17.	None	17.	None - pH 5.6	17.	1.26 M Sodium phosphate monobasic monohydrate
					0.14 M Potassium phosphate dibasic
18.	None	18.	None - pH 6.9	18.	0.49 M Sodium phosphate monobasic monohydrate
					0.91 M Potassium phosphate dibasic
19.	None	19.	None - pH 8.2	19.	0.056 M Sodium phosphate monobasic monohydrate
					1.344 M Potassium phosphate dibasic
20.	None	20.	0.1 M HEPES pH 7.5	20.	1.4 M Sodium citrate tribasic dihydrate
21.	None	21.	None	21.	1.8 M Ammonium citrate tribasic pH 7.0
22.	None	22.	None	22.	0.8 M Succinic acid pH 7.0
23.	None	23.	None	23.	2.1 M DL-Malic acid pH 7.0
24.	None	24.	None	24.	2.8 M Sodium acetate trihydrate pH 7.0
25.	None	25.	None	25.	3.5 M Sodium formate pH 7.0
26.	None	26.	None	26.	1.1 M Ammonium tartrate dibasic pH 7.0
27.	None	27.	None	27.	2.4 M Sodium malonate pH 7.0
28.	None	28.	None	28.	35% v/v Tacsimate™ pH 7.0
29.	None	29.	None	29.	60% v/v Tacsimate™ pH 7.0
30.	0.1 M Sodium chloride	30.	0.1 M BIS-TRIS pH 6.5	30.	1.5 M Ammonium sulfate
31.	0.8 M Potassium sodium tartrate tetrahydrate	31.	0.1 M Tris pH 8.5	31.	0.5% w/v Polyethylene glycol monomethyl ether 5,000
32.	1.0 M Ammonium sulfate	32.	0.1 M BIS-TRIS pH 5.5	32.	1% w/v Polyethylene glycol 3,350
33.	1.1 M Sodium malonate pH 7.0	33.	0.1 M HEPES pH 7.0	33.	0.5% w/v Jeffamine® ED-2001 pH 7.0
34.	1.0 M Succinic acid pH 7.0	34.	0.1 M HEPES pH 7.0	34.	1% w/v Polyethylene glycol monomethyl ether 2,000
35.	1.0 M Ammonium sulfate	35.	0.1 M HEPES pH 7.0	35.	0.5% w/v Polyethylene glycol 8,000
36.	15% w/v Tacsimate™ pH 7.0	36.	0.1 M HEPES pH 7.0	36.	2% w/v Polyethylene glycol 3,350
37.	None	37.	None	37.	25% w/v Polyethylene glycol 1,500
38.	None	38.	0.1 M HEPES pH 7.0	38.	30% v/v Jeffamine® M-600® pH 7.0
39.	None	39.	0.1 M HEPES pH 7.0	39.	30% v/v Jeffamine® ED-2001 pH 7.0
40.	None	40.	0.1 M Citric acid pH 3.5	40.	25% w/v Polyethylene glycol 3,350
41.	None	41.	0.1 M Sodium acetate trihydrate pH 4.5	41.	25% w/v Polyethylene glycol 3,350
42.	None	42.	0.1 M BIS-TRIS pH 5.5	42.	25% w/v Polyethylene glycol 3,350
43.	None	43.	0.1 M BIS-TRIS pH 6.5	43.	25% w/v Polyethylene glycol 3,350
44.	None	44.	0.1 M HEPES pH 7.5	44.	25% w/v Polyethylene glycol 3,350
45.	None	45.	0.1 M Tris pH 8.5	45.	25% w/v Polyethylene glycol 3,350
46.	None	46.	0.1 M BIS-TRIS pH 6.5	46.	20% w/v Polyethylene glycol monomethyl ether 5,000
47.	None	47.	0.1 M BIS-TRIS pH 6.5	47.	28% w/v Polyethylene glycol monomethyl ether 2,000
48.	0.2 M Calcium chloride dihydrate	48.	0.1 M BIS-TRIS pH 5.5	48.	45% v/v (+/-)-2-Methyl-2,4-pentanediol

◊ Buffer pH is that of a 1.0 M stock prior to dilution with other reagent components: pH with HCl or NaOH.

Index HT

Index HT™

HR2-134 Reagent Formulation

Well #	Salt	Well #	Buffer ◊	Well #	Precipitant
1. (A1)	None	1. (A1)	0.1 M Citric acid pH 3.5	1. (A1)	2.0 M Ammonium sulfate
2. (A2)	None	2. (A2)	0.1 M Sodium acetate trihydrate pH 4.5	2. (A2)	2.0 M Ammonium sulfate
3. (A3)	None	3. (A3)	0.1 M BIS-TRIS pH 5.5	3. (A3)	2.0 M Ammonium sulfate
4. (A4)	None	4. (A4)	0.1 M BIS-TRIS pH 6.5	4. (A4)	2.0 M Ammonium sulfate
5. (A5)	None	5. (A5)	0.1 M HEPES pH 7.5	5. (A5)	2.0 M Ammonium sulfate
6. (A6)	None	6. (A6)	0.1 M Tris pH 8.5	6. (A6)	2.0 M Ammonium sulfate
7. (A7)	None	7. (A7)	0.1 M Citric acid pH 3.5	7. (A7)	3.0 M Sodium chloride
8. (A8)	None	8. (A8)	0.1 M Sodium acetate trihydrate pH 4.5	8. (A8)	3.0 M Sodium chloride
9. (A9)	None	9. (A9)	0.1 M BIS-TRIS pH 5.5	9. (A9)	3.0 M Sodium chloride
10. (A10)	None	10. (A10)	0.1 M BIS-TRIS pH 6.5	10. (A10)	3.0 M Sodium chloride
11. (A11)	None	11. (A11)	0.1 M HEPES pH 7.5	11. (A11)	3.0 M Sodium chloride
12. (A12)	None	12. (A12)	0.1 M Tris pH 8.5	12. (A12)	3.0 M Sodium chloride
13. (B1)	None	13. (B1)	0.1 M BIS-TRIS pH 5.5	13. (B1)	0.3 M Magnesium formate dihydrate
14. (B2)	None	14. (B2)	0.1 M BIS-TRIS pH 6.5	14. (B2)	0.5 M Magnesium formate dihydrate
15. (B3)	None	15. (B3)	0.1 M HEPES pH 7.5	15. (B3)	0.5 M Magnesium formate dihydrate
16. (B4)	None	16. (B4)	0.1 M Tris pH 8.5	16. (B4)	0.3 M Magnesium formate dihydrate
17. (B5)	None	17. (B5)	None - pH 5.6	17. (B5)	1.26 M Sodium phosphate monobasic monohydrate, 0.14 M Potassium phosphate dibasic
18. (B6)	None	18. (B6)	None - pH 6.9	18. (B6)	0.49 M Sodium phosphate monobasic monohydrate, 0.91 M Potassium phosphate dibasic
19. (B7)	None	19. (B7)	None - pH 8.2	19. (B7)	0.056 M Sodium phosphate monobasic monohydrate, 1.344 M Potassium phosphate dibasic
20. (B8)	None	20. (B8)	0.1 M HEPES pH 7.5	20. (B8)	1.4 M Sodium citrate tribasic dihydrate
21. (B9)	None	21. (B9)	None	21. (B9)	1.8 M Ammonium citrate tribasic pH 7.0
22. (B10)	None	22. (B10)	None	22. (B10)	0.8 M Succinic acid pH 7.0
23. (B11)	None	23. (B11)	None	23. (B11)	2.1 M DL-Malic acid pH 7.0
24. (B12)	None	24. (B12)	None	24. (B12)	2.8 M Sodium acetate trihydrate pH 7.0
25. (C1)	None	25. (C1)	None	25. (C1)	3.5 M Sodium formate pH 7.0
26. (C2)	None	26. (C2)	None	26. (C2)	1.1 M Ammonium tartrate dibasic pH 7.0
27. (C3)	None	27. (C3)	None	27. (C3)	2.4 M Sodium malonate pH 7.0
28. (C4)	None	28. (C4)	None	28. (C4)	35% w/v Tacsimate pH 7.0
29. (C5)	None	29. (C5)	None	29. (C5)	60% w/v Tacsimate pH 7.0
30. (C6)	0.1 M Sodium chloride	30. (C6)	0.1 M BIS-TRIS pH 6.5	30. (C6)	1.5 M Ammonium sulfate
31. (C7)	0.8 M Potassium sodium tartrate tetrahydrate	31. (C7)	0.1 M Tris pH 8.5	31. (C7)	0.5% w/v Polyethylene glycol monomethyl ether 5,000
32. (C8)	1.0 M Ammonium sulfate	32. (C8)	0.1 M BIS-TRIS pH 5.5	32. (C8)	1% w/v Polyethylene glycol 3,350
33. (C9)	1.1 M Sodium malonate pH 7.0	33. (C9)	0.1 M HEPES pH 7.0	33. (C9)	0.5% w/v Jeffamine® ED-2001 pH 7.0
34. (C10)	1.0 M Succinic acid pH 7.0	34. (C10)	0.1 M HEPES pH 7.0	34. (C10)	1% w/v Polyethylene glycol monomethyl ether 2,000
35. (C11)	1.0 M Ammonium sulfate	35. (C11)	0.1 M HEPES pH 7.0	35. (C11)	0.5% w/v Polyethylene glycol 8,000
36. (C12)	15% w/v Tacsimate pH 7.0	36. (C12)	0.1 M HEPES pH 7.0	36. (C12)	2% w/v Polyethylene glycol 3,350
37. (D1)	None	37. (D1)	None	37. (D1)	25% w/v Polyethylene glycol 1,500
38. (D2)	None	38. (D2)	0.1 M HEPES pH 7.0	38. (D2)	30% w/v Jeffamine® M-600 pH 7.0
39. (D3)	None	39. (D3)	0.1 M HEPES pH 7.0	39. (D3)	30% w/v Jeffamine® ED-2001 pH 7.0
40. (D4)	None	40. (D4)	0.1 M Citric acid pH 3.5	40. (D4)	25% w/v Polyethylene glycol 3,350
41. (D5)	None	41. (D5)	0.1 M Sodium acetate trihydrate pH 4.5	41. (D5)	25% w/v Polyethylene glycol 3,350
42. (D6)	None	42. (D6)	0.1 M BIS-TRIS pH 5.5	42. (D6)	25% w/v Polyethylene glycol 3,350
43. (D7)	None	43. (D7)	0.1 M BIS-TRIS pH 6.5	43. (D7)	25% w/v Polyethylene glycol 3,350
44. (D8)	None	44. (D8)	0.1 M HEPES pH 7.5	44. (D8)	25% w/v Polyethylene glycol 3,350
45. (D9)	None	45. (D9)	0.1 M Tris pH 8.5	45. (D9)	25% w/v Polyethylene glycol 3,350
46. (D10)	None	46. (D10)	0.1 M BIS-TRIS pH 6.5	46. (D10)	20% w/v Polyethylene glycol monomethyl ether 5,000
47. (D11)	None	47. (D11)	0.1 M BIS-TRIS pH 6.5	47. (D11)	28% w/v Polyethylene glycol monomethyl ether 2,000
48. (D12)	0.2 M Calcium chloride dihydrate	48. (D12)	0.1 M BIS-TRIS pH 5.5	48. (D12)	45% w/v (+/-)-2-Methyl-2,4-pentanediol

◊ Buffer pH is that of a 1.0 M stock prior to dilution with other reagent components: pH with HCl or NaOH.

Hampton (HT) – Screen I

[A1] Tube : 1 0.02M Calcium chloride dihydrate 0.1M Sodium acetate trihydrate pH4.6 30%v/v 2-methyl-2,4-pentenediol	[A2] Tube : 2 0.4M Potassium sodium tartrate tetrahydrate	[A3] Tube : 3 0.4M Ammonium dihydrogen phosphate	[A4] Tube : 4 0.1M Tris hydrochloride pH8.5 2M Ammonium sulfate	[A5] Tube : 5 0.2M tri-Sodium citrate dihydrate 0.1M Sodium HEPES pH7.5 30%v/v 2-methyl-2,4-pentenediol	[A6] Tube : 6 0.2M magnesium chloride hexahydrate 0.1M Tris hydrochloride pH8.5 30%w/v polyethylene glycol 4000
[B1] Tube : 7 0.1M sodium cacodylate pH6.5 1.4M sodium acetate trihydrate	[B2] Tube : 8 0.2M tri-sodium citrate dihydrate 0.1M sodium cacodylate pH6.5 30%v/v iso-propanol	[B3] Tube : 9 0.2M ammonium acetate 0.1M tri-sodium citrate dihydrate pH5.6 30%w/v polyethylene glycol 4000	[B4] Tube : 10 0.2M ammonium acetate 0.1M Sodium acetate trihydrate pH4.6 30%w/v polyethylene glycol 4000	[B5] Tube : 11 0.1M tri-sodium citrate dihydrate pH5.6 1M Ammonium dihydrogen phosphate	[B6] Tube : 12 0.2M magnesium chloride hexahydrate 0.1M Sodium HEPES pH7.5 30%v/v iso-propanol
[C1] Tube : 13 0.2M tri-Sodium citrate dihydrate 0.1M Tris hydrochloride pH8.5 30%v/v polyethylene glycol 400	[C2] Tube : 14 0.2M Calcium chloride dihydrate 0.1M Sodium HEPES pH7.5 28%v/v polyethylene glycol 400	[C3] Tube : 15 0.2M ammonium sulfate 0.1M sodium cacodylate pH6.5 30%w/v polyethylene glycol 8000	[C4] Tube : 16 0.1M Sodium HEPES pH7.5 1.5M lithium sulfate monohydrate	[C5] Tube : 17 0.2M lithium sulfate monohydrate 0.1M Tris hydrochloride pH8.5 30%w/v polyethylene glycol 4000	[C6] Tube : 18 0.2M magnesium acetate tetrahydrate 0.1M sodium cacodylate pH6.5 20%w/v polyethylene glycol 8000
[D1] Tube : 19	[D2] Tube : 20	[D3] Tube : 21	[D4] Tube : 22	[D5] Tube : 23	[D6] Tube : 24

0.2M ammonium acetate 0.1M Tris hydrochloride pH8.5 30%v/v iso-propanol	0.2M ammonium sulfate 0.1M Sodium acetate trihydrate pH4.6 25%w/v polyethylene glycol 4000	0.2M magnesium acetate tetrahydrate 0.1M sodium cacodylate pH6.5 30%v/v 2-methyl-2,4-pentanediol	0.2M sodium acetate trihydrate 0.1M Tris hydrochloride pH8.5 30%w/v polyethylene glycol 4000	0.2M magnesium chloride hexahydrate 0.1M Sodium HEPES pH7.5 30%v/v polyethylene glycol 400	0.2M Calcium chloride dihydrate 0.1M Sodium acetate trihydrate pH4.6 20%v/v iso-propanol
[E1] Tube : 25 0.1M imidazole pH6.5 1M sodium acetate trihydrate	[E2] Tube : 26 0.2M ammonium acetate 0.1M tri-sodium citrate dihydrate pH5.6 30%v/v 2-methyl-2,4-pentanediol	[E3] Tube : 27 0.2M tri-Sodium citrate dihydrate 0.1M Sodium HEPES pH7.5 20%v/v iso-propanol	[E4] Tube : 28 0.2M sodium acetate trihydrate 0.1M sodium cacodylate pH6.5 30%w/v polyethylene glycol 8000	[E5] Tube : 29 0.1M Sodium HEPES pH7.5 0.8M Potassium sodium tartrate tetrahydrate	[E6] Tube : 30 0.2M ammonium sulfate 30%w/v polyethylene glycol 8000
[F1] Tube : 31 0.2M ammonium sulfate 30%w/v polyethylene glycol 4000	[F2] Tube : 32 2M Ammonium sulfate	[F3] Tube : 33 4M sodium formate	[F4] Tube : 34 0.1M sodium acetate trihydrate pH4.6 2M sodium formate	[F5] Tube : 35 0.1M Sodium HEPES pH7.5 0.8M sodium dihydrogen phosphate 0.8M potassium dihydrogen phosphate	[F6] Tube : 36 0.1M Tris hydrochloride pH8.5 8%w/v polyethylene glycol 8000
[G1] Tube : 37 0.1M Sodium acetate trihydrate pH4.6 8%w/v	[G2] Tube : 38 0.1M Sodium HEPES pH7.5 1.4M tri-	[G3] Tube : 39 0.1M Sodium HEPES pH7.5 2%v/v	[G4] Tube : 40 0.1M tri-sodium citrate dihydrate pH5.6	[G5] Tube : 41 0.1M Sodium HEPES pH7.5 10%v/v iso-propanol	[G6] Tube : 42 0.05M potassium dihydrogen phosphate 20%w/v polyethylene

polyethylene glycol 4000	sodium citrate dihydrate	polyethylene glycol 400 2M ammonium sulfate	20%v/v iso-propanol 20%w/v polyethylene glycol 4000	20%w/v polyethylene glycol 4000	glycol 8000
[H1] Tube : 43 30%w/v polyethylene glycol 1500	[H2] Tube : 44 0.2M magnesium formate	[H3] Tube : 45 0.2M zinc acetate dihydrate 0.1M sodium cacodylate pH6.5 18%w/v polyethylene glycol 8000	[H4] Tube : 46 0.2M calcium acetate hydrate 0.1M sodium cacodylate pH6.5 18%w/v polyethylene glycol 8000	[H5] Tube : 47 0.1M Sodium acetate trihydrate pH4.6 2M Ammonium sulfate	[H6] Tube : 48 0.1M Tris hydrochloride pH8.5 2M Ammonium dihydrogen phosphate
[I1] Tube : 49 1M lithium sulfate monohydrate 2%w/v polyethylene glycol 8000	[I2] Tube : 50 0.5M lithium sulfate monohydrate 15%w/v polyethylene glycol 8000				

Hampton (HT) – Screen I

[A1] Tube : 1 2.0M sodium chloride 10% PEG 6000	[A2] Tube : 2 0.5M sodium chloride 0.01M hexadecyltrimethylammonium bromide 0.01M magnesium chloride hexahydrate	[A3] Tube : 3 25%v/v ethylene glycol	[A4] Tube : 4 35%v/v dioxane	[A5] Tube : 5 2.0M ammonium sulfate 5%v/v iso-propanol	[A6] Tube : 6 pH7.0 1M imidazole
[B1] Tube : 7 10%w/v polyethylene glycol	[B2] Tube : 8 1.5M sodium chloride 10%v/v ethanol	[B3] Tube : 9 0.1M sodium	[B4] Tube : 10 0.2M sodium chloride	[B5] Tube : 11 0.01M cobaltous chloride	[B6] Tube : 12 0.1M cadmium chloride

ne glycol 1000 10%w/v polyethyle ne glycol 8000		acetate trihydrate pH4.6 2M sodium chloride	0.1M sodium acetate trihydrate pH4.6 30%v/v MPD	hexahydra te 0.1M sodium acetate trihydrate pH4.6 1M 1,6 hexanediol	dihydrate 0.1M sodium acetate trihydrate pH4.6 30%v/v polyethyle ne glycol 400
[C1] Tube : 13 0.2M ammoniu m sulfate 0.1M sodium acetate trihydrate pH4.6 30%w/v polyethyle ne glycol monomet hyl ether 2000	[C2] Tube : 14 0.2M potassium sodium tartrate tetrahydrate 0.1M tri-sodium citrate dihydrate pH5.6 2M ammonium sulfate	[C3] Tube : 15 0.5M ammoniu m sulfate 0.1M tri- sodium citrate dihydrate pH5.6 1M lithium sulfate monohydr ate	[C4] Tube : 16 0.5M sodium chloride 0.1M tri- sodium citrate dihydrate pH5.6 2%v/v ethylene imine polymer	[C5] Tube : 17 0.1M tri- sodium citrate dihydrate pH5.6 35%v/v tert- butanol	[C6] Tube : 18 0.01M ferric chloride hexahydr ate 0.1M tri- sodium citrate dihydrate pH5.6 10%v/v jeffamine M-600
[D1] Tube : 19 0.1M tri- sodium citrate dihydrate pH5.6 2.5M 1,6 hexanedio l	[D2] Tube : 20 0.1M MES pH6.5 1.6M magnesium sulfate heptahydrate	[D3] Tube : 21 0.1M sodium dihydroge n phosphate 0.1M MES pH6.5 2M sodium chloride 0.1 M potassium dihydroge n phosphate	[D4] Tube : 22 0.1M MES pH6.5 12%w/v polyethyle ne glycol 20,000	[D5] Tube : 23 1.6M ammoniu m sulfate 0.1M MES pH6.5 10%v/v dioxane	[D6] Tube : 24 0.05M cesium chloride 0.1M MES pH6.5 30%v/v jeffamine M-600
[E1] Tube : 25 0.01M	[E2] Tube : 26 0.2M ammonium sulfate 0.1M MES pH6.5	[E3] Tube : 27 0.01M zinc	[E4] Tube : 28	[E5] Tube : 29 0.5M	[E6] Tube : 30

cobaltous chloride hexahydr ate 0.1M MES pH6.5 1.8M ammoniu m sulfate	30%w/v polyethylene glycol monomethyl ether 5000	sulfate heptahydr ate 0.1M MES pH6.5 25%v/v polyethyle ne glycol monometh yl ether 550	pH6.5 1.6M tri- sodium citrate dihydrate	ammoniu m sulfate 0.1M HEPES pH7.5 30%v/v MPD	0.1M HEPES pH7.5 10%w/v polyethyle ne glycol 6000 5%v/v MPD
[F1] Tube : 31 0.1M HEPES pH7.5 20%v/v jeffamine M-600	[F2] Tube : 32 0.1M sodium chloride 0.1M HEPES pH7.5 1.6M ammonium sulfate	[F3] Tube : 33 0.1M HEPES pH7.5 2M ammoniu m formate	[F4] Tube : 34 0.05M cadmium sulfate hydrate 0.1M HEPES pH7.5 1M sodium acetate trihydrate	[F5] Tube : 35 0.1M HEPES pH7.5 70%v/v MPD	[F6] Tube : 36 0.1M HEPES pH7.5 4.3M sodium chloride
[G1] Tube : 37 0.1M HEPES pH7.5 10%w/v polyethyle ne glycol 8000	[G2] Tube : 38 0.1M HEPES pH7.5 20%w/v polyethylene glycol 10,000 8%v/v ethylene glycol	[G3] Tube : 39 0.2M magnesi um chloride hexahydra te 0.1M tris pH8.5 3.4M 1,6 hexanediol	[G4] Tube : 40 0.1M tris pH8.5 25%v/v tert- butanol	[G5] Tube : 41 0.01M nickel (II) chloride hexahydra te 0.1M tris pH8.5 1M lithium sulfate monohydr ate	[G6] Tube : 42 1.5M ammoniu m sulfate 0.1M tris pH8.5 12%v/v glycerol anhydrou s
[H1] Tube : 43 0.2M ammoniu m dihydroge n phosphate	[H2] Tube : 44 0.1M tris pH8.5 20%v/v ethanol	[H3] Tube : 45 0.01M nickel (II) chloride hexahydra te 0.1M tris	[H4] Tube : 46 0.1M sodium chloride 0.1M bicine pH9.0	[H5] Tube : 47 0.1M bicine pH9.0 2M magnesi um	[H6] Tube : 48 0.1M bicine pH9.0 2%v/v dioxane

0.1M tris pH8.5 50%v/v MPD		pH8.5 20%w/v polyethyle ne glycol monomet hyl ether 2000	20%v/v polyethyle ne glycol monomet hyl ether 550	m chloride hexahydra te	10%w/v polyethyle ne glycol 20,000
-------------------------------------	--	--	--	-------------------------------	---

* Newcastle screen formulations not available at time of writing.

Appendix IV: DNA sequencing outputs

SEQ2190

```
Query:      114 ATATGGATATGAAAAGCGATGCCAAAAAGGTATTAGCAACATTGCAAAAAGGTAATCATA 173
             ||  |||
Sbjct: 2203279 ATGGGGATATGAAAAGCGATGCCAAAAAGGTATTAGCAACATTGCAAAAAGGTAATCATA 2203338

Query:      174 AACAAAGAGTTAAAAAAGTTCTTAGAACTATTGGTGATGATAGTTTGCTATTATTATGGT 233
             |||
Sbjct: 2203339 AACAAAGAGTTAAAAAAGTTCTTAGAACTATTGGTGATGATAGTTTGCTATTATTATGGT 2203398

Query:      234 TACTTGTATCGGGTGATTCCAATTTAGATGAAAGTAATAAAAGTTTGTATTATTATCAA 293
             |||
Sbjct: 2203399 TACTTGTATCGGGTGATTCCAATTTAGATGAAAGTAATAAAAGTTTGTATTATTATCAA 2203458

Query:      294 TTTGGAATACTGGATCTAGAGTTGGTAAAGTGATATCTCCTACTGAATACTTTAAACAAG 353
             |||
Sbjct: 2203459 TTTGGAATACTGGATCTAGAGTTGGTAAAGTGATATCTCCTACTGAATACTTTAAACAAG 2203518

Query:      354 AATCAAAAAACACTGAAAATAGGCAAGCTGTTTATAGTGAATTTAAACAACGCGTAGAAG 413
             |||
Sbjct: 2203519 AATCAAAAAACACTGAAAATAGGCAAGCTGTTTATAGTGAATTTAAACAACGCGTAGAAG 2203578

Query:      414 AACGATCTAAAAAGCCAAGGAAGCCACTGAAGCACTAAACCAAAAAGCTCAATTGGAGG 473
             |||
Sbjct: 2203579 AACGATCTAAAAAGCCAAGGAAGCCACTGAAGCACTAAACCAAAAAGCTCAATTGGAGG 2203638

Query:      474 CTACAGTTAAAAATATAAACGAAGAACTTGAAAAAACTAGAGAAGGTTTTAAAGTTGTGA 533
             |||
Sbjct: 2203639 CTACAGTTAAAAATATAAACGAAGAACTTGAAAAAACTAGAGAAGGTTTTAAAGTTGTGA 2203698

Query:      534 GTGAAAACTCAACAAATTAGAAAAACAATTAATGGCTGAGAAAAATAAAACACGTACAG 593
             |||
Sbjct: 2203699 GTGAAAACTCAACAAATTAGAAAAACAATTAATGGCTGAGAAAAATAAAACACGTACAG 2203758

Query:      594 CTGAAGAACTGCCAAACAAGCTAAAACCGATAAAGAAAGAGCAGAGGCTGAAGCTAAAA 653
             |||
Sbjct: 2203759 CTGAAGAACTGCCAAACAAGCTAAAACCGATAAAGAAAGAGCAGAGGCTGAAGCTAAAA 2203818

Query:      654 AAGCGAAAGAGGAAGCTAAGACTGCGGAAGGAAAAGTTAAACAAGCAGAACTGAGAAAA 713
             |||
Sbjct: 2203819 AAGCGAAAGAGGAAGCTAAGACTGCGGAAGGAAAAGTTAAACAAGCAGAACTGAGAAAA 2203878

Query:      714 GGAATGCGGAAGCTAAGGCTAGAACCGCTGAAGAAGAAGCTAAACAAGCTACGGCTGATA 773
             |||
Sbjct: 2203879 GGAATGCGGAAGCTAAGGCTAGAACCGCTGAAGAAGAAGCTAAACAAGCTACGGCTGATA 2203938
```

Query: 774 AAGAAAAGGCAGAGACTGAAGCTAAAAAGCGAAAGAAGAAGCTAAGACCGCTAAAGAAG 833
 |||
 Sbjct: 2203939 AAGAAAAGGCAGAGACTGAAGCTAAAAAGCGAAAGAAGAAGCTAAGACCGCTAAAGAAG 2203998

Query: 834 CTGCTCATCAAGAACAAGAAAAAGCTAAACAATTAGAACAAGCTAATCAACAGGCTAACC 893
 |||
 Sbjct: 2203999 CTGCTCATCAAGAACAAGAAAAAGCTAAACAATTAGAACAAGCTAATCAACAGGCTAACC 2204058

Query: 894 AAAGAGCTAATCTGGNNCTTG 953
 ||| |
 Sbjct: 2204059 AAAGAGCTAATCTGGCAGAAAAATCTAAAAAGATTTAGAAACTCAAAAAGAAAACTTG 2204118

Query: 954 AACAAAGAGATTTAAGA 969
 |||
 Sbjct: 2204119 AACAAAGAGATTAAAGA 2204134

SEQ2180

Query: 118 GATATGAGTGCGGACGTNNNNNNNNNTTGAGCAAGTTTACAAATCAGATAGTGGTAAGTTA 177
 ||| |
 Sbjct: 2192072 GATATGAGTGCGGACGTAAAAAAATTGAGCAAGTTTACAAATCAGATAGTGGTAAGTTA 2192131

Query: 178 AAAACAGTATTAGGACAAATTAGTGATCCCAATACATTGTTTGCTTTATTTGCTATTATA 237
 |||
 Sbjct: 2192132 AAAACAGTATTAGGACAAATTAGTGATCCCAATACATTGTTTGCTTTATTTGCTATTATA 2192191

Query: 238 TCAGCCGATGTAAATATAGATCCAGAAAGTTTTAGTTTTATATTAAATGGCCCAATCACA 297
 |||
 Sbjct: 2192192 TCAGCCGATGTAAATATAGATCCAGAAAGTTTTAGTTTTATATTAAATGGCCCAATCACA 2192251

Query: 298 CCATTATTATCATATTTCAACAATCCAAACAAAAATGATAATAGAAAGGCAGTTGTTAAT 357
 |||
 Sbjct: 2192252 CCATTATTATCATATTTCAACAATCCAAACAAAAATGATAATAGAAAGGCAGTTGTTAAT 2192311

Query: 358 GAAATNNNNNNNNGAGTCTCAGCTTTGGAGCAAGCTCAAAAATTAGAAACTAAAGCTAAA 417
 ||| |
 Sbjct: 2192312 GAAATAAAAAAAGAGTCTCAGCTTTGGAGCAAGCTCAAAAATTAGAAACTAAAGCTAAA 2192371

Query: 418 ATAGAGAAACAGATTGCAGAAAGTGAATTAGAAAAAACTAGGAATGCTTTCAAAGTAACA 477
 |||
 Sbjct: 2192372 ATAGAGAAACAGATTGCAGAAAGTGAATTAGAAAAAACTAGGAATGCTTTCAAAGTAACA 2192431

Query: 478 ATTGAAAATTTACATAAACTAGAGGGATTATTAGACACTGAAAAACAAAAAACTCGAAAA 537
 |||
 Sbjct: 2192432 ATTGAAAATTTACATAAACTAGAGGGATTATTAGACACTGAAAAACAAAAAACTCGAAAA 2192491

Query: 538 GTTGAGGAAGATTACCAACAAGCTAAAACCGATAAGAAAAAGCAGAGGCTGATAAAAGG 597
 |||

Sbjct: 2192492 GTTGAGGAAGATTACCAACAAGCTAAAACCGATAAAAGAAAAGCAGAGGCTGATAAAAGG 2192551

Query: 598 AATGCGGAAACTAAGGCTAGAACCGCTGAAGAAGAAGCTAAACAAGCTACGGCTGATAAA 657
 |||

Sbjct: 2192552 AATGCGGAAACTAAGGCTAGAACCGCTGAAGAAGAAGCTAAACAAGCTACGGCTGATAAA 2192611

Query: 658 GAAAAGGCAGAGACTGAAGCTAAAAAGCGAAAGAAGAAGCTAAGACCGCTAAAGAAGCT 717
 |||

Sbjct: 2192612 GAAAAGGCAGAGACTGAAGCTAAAAAGCGAAAGAAGAAGCTAAGACCGCTAAAGAAGCT 2192671

Query: 718 GCTCATCAAGAACAAGAAAAAGCTAAACAATTAGAACAAGCTAATCAACAGGCTAACCAA 777
 |||

Sbjct: 2192672 GCTCATCAAGAACAAGAAAAAGCTAAACAATTAGAACAAGCTAATCAACAGGCTAACCAA 2192731

Query: 778 AGAGCTAATCTGGNNCTTGAA 837
 |||

Sbjct: 2192732 AGAGCTAATCTGGCAGAAAAATCTAAAAAAGATTAGAAACTCAAAAAGAAAAACTTGAA 2192791

Query: 838 CAAGAGATTAAAGAAGCTACAGAAGCTAAAAACAAAGCCGAGCAAAAATTAAAAGACCTA 897
 |||

Sbjct: 2192792 CAAGAGATTAAAGAAGCTACAGAAGCTAAAAACAAAGCCGAGCAAAAATTAAAAGACCTA 2192851

Query: 898 CAAGATTCAGCTAGTCAAGGCAGTGAATTGTCTAAACAGCTGTAAAAGAATAAGAATGA 957
 |||

Sbjct: 2192852 CAAGATTCAGCTAGTCAAGGCAGTGAATTGTCTAAACAGCTGTAAAAGAAAAAGAA-GA 2192910

Query: 958 CTTATCAACATATACTTCAAGAGCT 982
 |||

Sbjct: 2192911 ATTAACAACAAA-ACTTCAAGAGCT 2192934

SlaA (seq0849)

Query: 125 GAAGGGATAAATGATAAAATAGAAAATGGCACTGAAAGCGATATTAGCTTCCAAAATGGT 184
 |||

Sbjct: 824165 GAAGGGATAAATGATAAAATAGAAAATGGCACTGAAAGCGATATTAGCTTCCAAAATGGT 824224

Query: 185 GAACTCCTAAAAAATTATCTTATCCTAGAAAGGTGAACGAGTATACTTTGATTATGATAGA 244
 |||

Sbjct: 824225 GAACTCCTAAAAAATTATCTTATCCTAGAAAGGTGAACGAGTATACTTTGATTATGATAGA 824284

Query: 245 GCAACTCAAGATAAAGTATCAGATGATGTTCTAGAGATGGGAATGTTAGTTGAAGCTATA 304
 |||

Sbjct: 824285 GCAACTCAAGATAAAGTATCAGATGATGTTCTAGAGATGGGAATGTTAGTTGAAGCTATA 824344

Query: 305 AGTAAGGATTATTCTGAGAAGACATTCACCCCAGATAAAATATTTTAAAGCTAGTTGGCCT 364
 |||

Sbjct: 824345 AGTAAGGATTATTCTGAGAAGACATTCACCCCAGATAAAATATTTTAAAGCTAGTTGGCCT 824404

Query: 365 ATCCATGGTAACTATTGTGGACCAAGTCATAATGGGAATAACTTTACTTTGCCAGTAGTA 423
 |||
 Sbjct: 824405 ATCCATGGTAACTATTGTGGACCAAGTCATAATGGGAATAACTTTACTTTGCCAGTAGTA 824464

Query: 424 TATGTTTTGGATCAAGGTTGCCAAAACACGATAGTTGCTATAAGTGGGGTGCCAGTATT 483
 |||
 Sbjct: 824465 TATGTTTTGGATCAAGGTTGCCAAAACACGATAGTTGCTATAAGTGGGGTGCCAGTATT 824524

Query: 484 GGTGCTAATTGTGAATGTAACCGTCAGCTCGTTAATTACATAAAAGTTTATAGACGATGG 543
 |||
 Sbjct: 824525 GGTGCTAATTGTGAATGTAACCGTCAGCTCGTTAATTACATAAAAGTTTATAGACGATGG 824584

Query: 544 ATGCCAGCGAATGTCCTTGGCGTAGCCGATGCTATTAGAGTATATTTTGAGACAGTAGGT 603
 |||
 Sbjct: 824585 ATGCCAGCGAATGTCCTTGGCGTAGCCGATGCTATTAGAGTATATTTTGAGACAGTAGGT 824644

Query: 604 TCTATAGGATGTTAAATTTNNNNNNNNCGTTAGTGATTATTCGAAGTACTATATTTTGA 663
 |||
 Sbjct: 824645 TCTATAGGATGTTAAATTTCAAAAAAACGTTAGTGATTATTCGAAGTACTATATTTTGA 824704

Query: 664 GTATAGTGTTATTATGGATAGGAATGTTCTTTGTTTCCTTGGGATAAGACTCTATTCGGTG 723
 |||
 Sbjct: 824705 GTATAGTGTTATTATGGATAGGAATGTTCTTTGTTTCCTTGGGATAAGACTCTATTCGGTG 824764

Query: 724 TTTCGATTTATCTGTTTGTATGCAAAAGTTATTTATAGTTGTTTGGTTTGACGGCTATC 783
 |||
 Sbjct: 824765 TTTCGATTTATCTGTTTGTATGCAAAAGTTATTTATAGT-GTTTGGTTTGACGGCTATC 824823

Query: 784 TTCTTGCCCTTGGTGTCCAACCTGATAAATCTATTTATTCGCGGAATATTATTTGTTCTC 843
 ||| || ||| |||| || |||
 Sbjct: 824824 TTCATGGCCTTTGTGTCTAAAAAATAAATCTATTTATTTGGGGAATATTATTTATTTTC 824883

Query: 844 G 844
 |
 Sbjct: 824884 G 824884

SlaB (seq2155)

Query: 117 GAAGGGATAAATGATAAAATAGAAAATGGCACTGAAAGCGATATTAGCTTCCAAAATGGT 176
 |||
 Sbjct: 824165 GAAGGGATAAATGATAAAATAGAAAATGGCACTGAAAGCGATATTAGCTTCCAAAATGGT 824224

Query: 177 GAACTCCTAAAAAATTATCTTATCCTAGAAAGGTGAACGAGTATACTTTGATTATGATAGA 236
 |||
 Sbjct: 824225 GAACTCCTAAAAAATTATCTTATCCTAGAAAGGTGAACGAGTATACTTTGATTATGATAGA 824284

Query: 237 GCAACTCAAGATAAAGTATCAGATGATGTTCTAGAGATGGGAATGTTAGTTGAAGCTATA 296

|||||
Sbjct: 824285 GCAACTCAAGATAAAGTATCAGATGATGTTCTAGAGATGGGAATGTTAGTTGAAGCTATA 824344

Query: 297 AGTAAGGATTATTCTGAGAAGACATTCACCCCAGATAAAATATTTTAAAGCTAGTTGGCCT 356
|||||

Sbjct: 824345 AGTAAGGATTATTCTGAGAAGACATTCACCCCAGATAAAATATTTTAAAGCTAGTTGGCCT 824404

Query: 357 ATCCATGGTAACTATTGTGGACCAGGTCATAATGGGAATAACTTTACGTTGCCAGTAGTA 416
|||||

Sbjct: 824405 ATCCATGGTAACTATTGTGGACCAGGTCATAATGGGAATAACTTTACGTTGCCAGTAGTA 824464

Query: 417 GATGTTTTGGATCAAGGTTGCCAAAACACGATAGTTGCTATAAGTGGGGTGCCGGTATT 476
|||||

Sbjct: 824465 GATGTTTTGGATCAAGGTTGCCAAAACACGATAGTTGCTATAAGTGGGGTGCCGGTATT 824524

Query: 477 GGTGCTAATTGTGAATGTAACCGTCAGCTCGTTAATTACATAAAAAGTTTATAGACGATGG 536
|||||

Sbjct: 824525 GGTGCTAATTGTGAATGTAACCGTCAGCTCGTTAATTACATAAAAAGTTTATAGACGATGG 824584

Query: 537 ATGCCAGCGAATGTCCTTGGCGTAGCCGATGCTATTAGAGTATATTTTGAGACAGTAGGT 596
|||||

Sbjct: 824585 ATGCCAGCGAATGTCCTTGGCGTAGCCGATGCTATTAGAGTATATTTTGAGACAGTAGGT 824644

Query: 597 TCTATAGGATGTTAAATTTNNNNNNNNCGTTAGTGATTATTCGAAGTACTATATTTTGA 656
|||||

Sbjct: 824645 TCTATAGGATGTTAAATTTCAAAAAAACGTTAGTGATTATTCGAAGTACTATATTTTGA 824704

Query: 657 GTATAGTGTTATTATGGATAGGAATGTTCTTTGTTTCCTTGGGATAAGACTCTATTCGGTG 716
|||||

Sbjct: 824705 GTATAGTGTTATTATGGATAGGAATGTTCTTTGTTTCCTTGGGATAAGACTCTATTCGGTG 824764

Query: 717 TTTCGATTTATCTGTTTGTATGCAAAAGTTATTTATAGTGTTTGGTTTGACGGCTATCT 776
|||||

Sbjct: 824765 TTTCGATTTATCTGTTTGTATGCAAAAGTTATTTATAGTGTTTGGTTTGACGGCTATCT 824824

Query: 777 TCATGGCCTTTGTGTCTAAAAAATAAATCTATTTATTTGGGAATATTATTTATTTCTCG 836
|||||

Sbjct: 824825 TCATGGCCTTTGTGTCTAAAAAATAAATCTATTTATTTGGGAATATTATTTATTTCTCG 824884

SeLppC (seq0346)

Query: 834 ACGCCAAATGGAGAGAAAAAGCAGCTCCTAGTCAAGACAAGGATAAGCAAGAAAAGGTT 775
|||||

Sbjct: 328189 ACGACAAATGGAGAGAAAAAGCAGCTCCTAGTCAAGACAAGGATAAGCAAGAAAAGGTT 328248

Query: 774 GTCCGTCTAACCAATGATCAGCTAAGGGCTAGGGAAAATACCATGGCCACACTCTGGTAT 715
|||||

Sbjct: 328249 GTCCGTCTAACCAATGATCAGCTAAGGGCTAGGGAAAATACCATGGCCACACTCTGGTAT 328308

Query: 714 CAGCAGTCTGAAGAGGCTAAGGCTCTCTACTTACAGGGGTACCAAGTGGCTAAGCAGCGC 655
 |||

Sbjct: 328309 CAGCAGTCTGAAGAGGCTAAGGCTCTCTACTTACAGGGGTACCAAGTGGCTAAGCAGCGC 328368

Query: 654 TTAGATACATTATTGAGTCAAGCAACTGATAAACCATACTCTATTGTTTTGGACATTGAT 595
 |||

Sbjct: 328369 TTAGATACATTATTGAGTCAAGCAACTGATAAACCATACTCTATTGTTTTGGACATTGAT 328428

Query: 594 GAGACAGTGCTCGATAATAGTCCTTACCAGGCTAAAAATATTAAAGAGGGCACAGGCTTT 535
 |||

Sbjct: 328429 GAGACAGTGCTCGATAATAGTCCTTACCAGGCTAAAAATATTAAAGAGGGCACAGGCTTT 328488

Query: 534 ACGCCGGATTCTTGGGATAAATGGGTCCAAAAGAAATCAGCCAAGGCAGTAGCAGGTGCT 475
 |||

Sbjct: 328489 ACGCCGGATTCTTGGGATAAATGGGTCCAAAAGAAATCAGCCAAGGCAGTAGCAGGTGCT 328548

Query: 474 AAGGACTTCTTGCAAGTATGCAAATGACAAGGGCGTTCAGATTTACTATGTTTCTGACAGA 415
 |||

Sbjct: 328549 AAGGACTTCTTGCAAGTATGCAAATGACAAGGGCGTTCAGATTTACTATGTTTCTGACAGA 328608

Query: 414 ACGACTAAGCAAGTAGAGCCGACTATGGAAAATCTTGAAAAAGAAGGTATTCCAGTGCAA 355
 |||

Sbjct: 328609 ACGACTAAGCAAGTAGAGCCGACTATGGAAAATCTTGAAAAAGAAGGTATTCCAGTGCAA 328668

Query: 354 GGCAAGGATCATTCTTGTCTTAGAGGAGGGTGTAATAAATCAAAGAGGGCCGTCGTCAA 295
 |||

Sbjct: 328669 GGCAAGGATCATTCTTGTCTTAGAGGAGGGTGTAATAAATCAAAGAGGGCCGTCGTCAA 328728

Query: 294 AAGGTGCAAGAAACAACAATCTAGTGTGTTATTTGGTGACAATTTGCTTGATTTTGCA 235
 |||

Sbjct: 328729 AAGGTGCAAGAAACAACAATCTAGTGTGTTATTTGGTGACAATTTGCTTGATTTTGCA 328788

Query: 234 GAGTTTTCTAAGACCTCTCATGAAGATAGAAGAAAGCTTTTGGATCAGCTGCATGCAGAG 175
 |||

Sbjct: 328789 GAGTTTTCTAAGACCTCTCATGAAGATAGAAGAAAGCTTTTGGATCAGCTGCATGCAGAG 328848

Query: 174 TTTGGCAGCAAGTTTATCATCTTCCCTAATCCAATGTATGGATCATGGGAAAGCAGTGTC 115
 |||

Sbjct: 328849 TTTGGCAGCAAGTTTATCATCTTCCCTAATCCAATGTATGGATCATGGGAAAGCAGTGTC 328908

Query: 114 TATAATGGTCAAAAATTAGACAGTGAAGGCCAAATCAAAGCACGTGACAAGGCTNNAGGG 55
 |||

Sbjct: 328909 TATAATGGTCAAAAATTAGACAGTGAAGGCCAAATCAAAGCACGTGACAAGGCTTTAGAG 328968

Query: 54 GCTTATTAA 46
 |||

Sbjct: 328969 GCTTATTAA 328977

SzLppC (szo16870)

Query: 944 TTAATAAGCCTCTAAAGCTTTGTACGTGCTTTGATTTGGCCTTCACTGTCTAATTTTGG 885
|||||
Sbjct: 1886324 TTAATAAGCCTCTAAAGCTTTGTACGTGCTTTGATTTGGCCTTCACTGTCTAATTTTGG 1886383

Query: 884 ACCATTATAGACACTGCTTTCCCATGATCCATACATTGGATTAGGGAAGATGATAAACTT 825
|||||
Sbjct: 1886384 ACCATTATAGACACTGCTTTCCCATGATCCATACATTGGATTAGGGAAGATGATAAACTT 1886443

Query: 824 GCTGCCAAATTCTGCATGCAGCTGATCCAAAAGCTTTCTTCTATCTTCATGAGAGGTCTT 765
|||||
Sbjct: 1886444 GCTGCCAAATTCTGCATGCAGCTGATCCAAAAGCTTTCTTCTATCTTCATGAGAGGTCTT 1886503

Query: 764 AGAAAACCTCTGCAAAATCAAGCAAATTATCACCAAATAACAACACTAGATTTGTTGTTTC 705
|||||
Sbjct: 1886504 AGAAAACCTCTGCAAAATCAAGCAAATTATCACCAAATAACAACACTAGATTTGTTGTTTC 1886563

Query: 704 TTGCACTTTTTGACGACGGCCCTCTTTTGATTTTACACCCTCCTCTAAGAACAAGAAATG 645
|||||
Sbjct: 1886564 TTGCACTTTTTGACGACGGCCCTCTTTTGATTTTACACCCTCCTCTAAGAACAAGAAATG 1886623

Query: 644 ATCCTTGCCCTTGCACTGGAATACCTTCTTTTTCAAGATTTTCCATAGTCGGCTCTACTTG 585
|||||
Sbjct: 1886624 ATCCTTGCCCTTGCACTGGAATACCTTCTTTTTCAAGATTTTCCATAGTCGGCTCTACTTG 1886683

Query: 584 CTTAGTCGTTCTGTCAGAAACATAGTAAATCTGAACGCCCTTGTCATTTGCATACTGCAA 525
|||||
Sbjct: 1886684 CTTAGTCGTTCTGTCAGAAACATAGTAAATCTGAACGCCCTTGTCATTTGCATACTGCAA 1886743

Query: 524 GAAGTCCTTAGCACCTGCTACTGCCTTGGCTGATTTCTTTTGGACCCATTTATCCCAAGA 465
|||||
Sbjct: 1886744 GAAGTCCTTAGCACCTGCTACTGCCTTGGCTGATTTCTTTTGGACCCATTTATCCCAAGA 1886803

Query: 464 ATCCGGCGTAAAGCCTGTGCCCTCTTTAATATTTTGTAGCTTGGTAAGGACTATTATCGAG 405
|||||
Sbjct: 1886804 ATCCGGCGTAAAGCCTGTGCCCTCTTTAATATTTTGTAGCTTGGTAAGGACTATTATCGAG 1886863

Query: 404 CACTGTCTCATCAATGTCCAAAACAATAGAGTATGGTTTATCAGTTGCTTGACTCAATAA 345
|||||
Sbjct: 1886864 CACTGTCTCATCAATGTCCAAAACAATAGAGTATGGTTTATCAGTTGCTTGACTCAATAA 1886923

Query: 344 TGTATCTAAGCGCTGCTTAGCCACTTGGTACCCCTGTAGGTAGAGAGCCTTAGCCTCTTC 285
|||||
Sbjct: 1886924 TGTATCTAAGCGCTGCTTAGCCACTTGGTACCCCTGTAGGTAGAGAGCCTTAGCCTCTTC 1886983

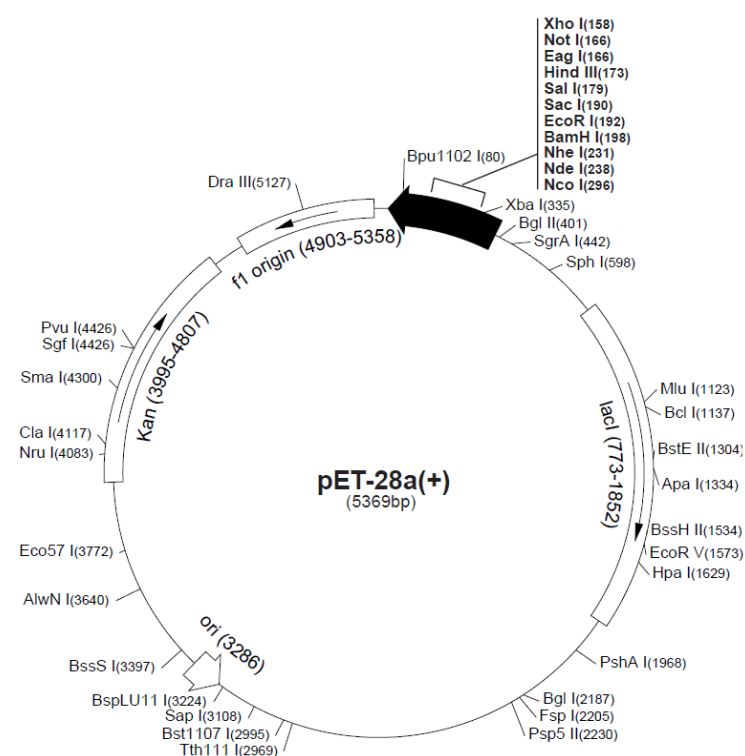
Query: 284 AGACTGCTGATACCAGAGTGTGGCCATAGTATTTCCCTAGCCCTTAGCTGATCATTTGGT 225
|||||
Sbjct: 1886984 AGACTGCTGATACCAGAGTGTGGCCATAGTATTTCCCTAGCCCTTAGCTGATCATTTGGT 1887043

```
Query:      224 TAGACGGACAACCTTTTCTTGCTTATCCTTGTCTTGACTAGGAGCTGCTTTTTTCTCTCC 165
            |||
Sbjct: 1887044 TAGACGGACAACCTTTTCTTGCTTATCCTTGTCTTGACTAGGAGCTGCTTTTTTCTCTCC 1887103

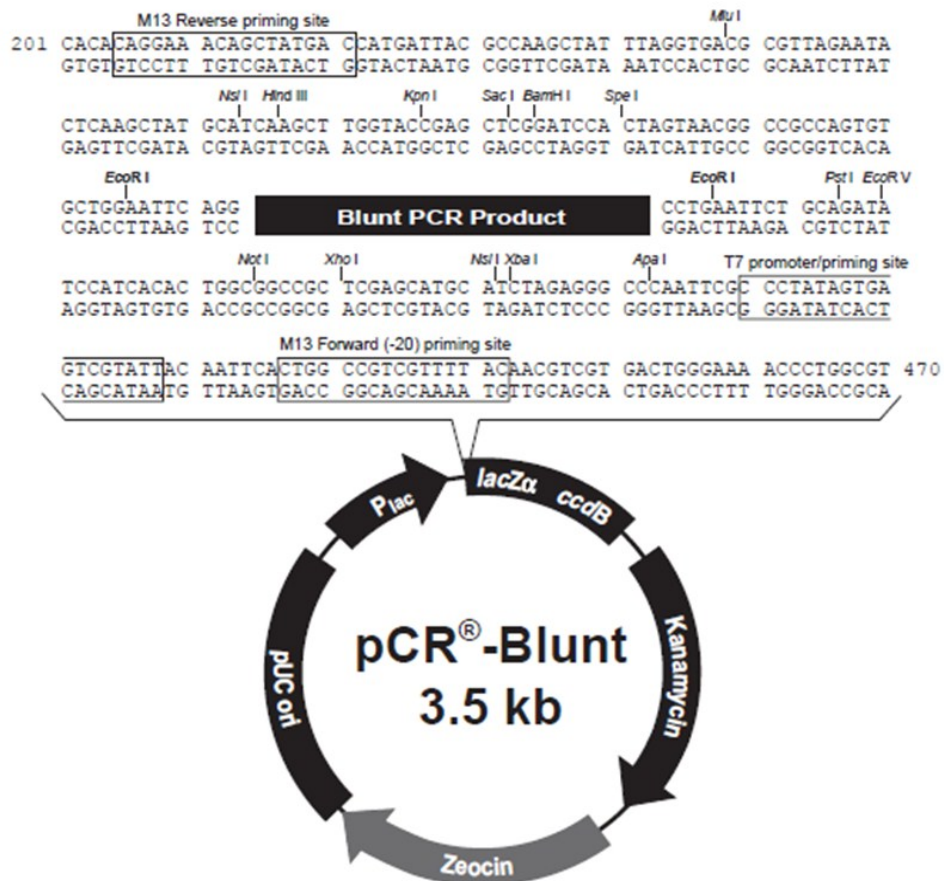
Query:      164 ATTTGTCGC 156
            |||
Sbjct: 1887104 ATTTGTCGC 1887112
```

–Appendix V: Cloning vector maps, molecular weight markers and standards

pET-28a expression vector

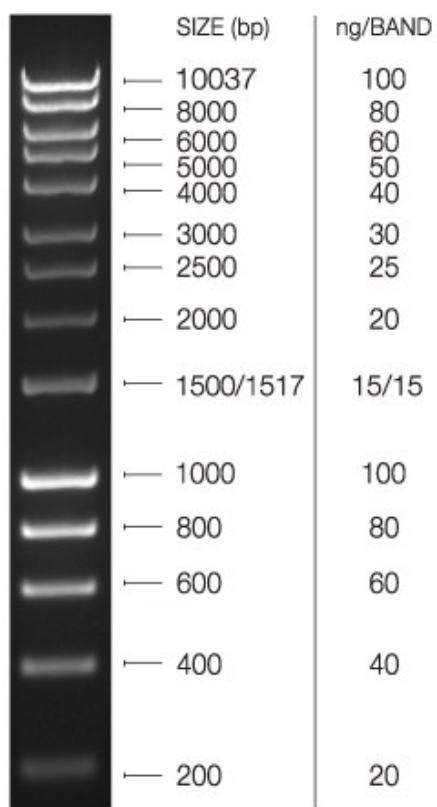


pCR[®]-Blunt cloning vector.



DNA electrophoresis standards

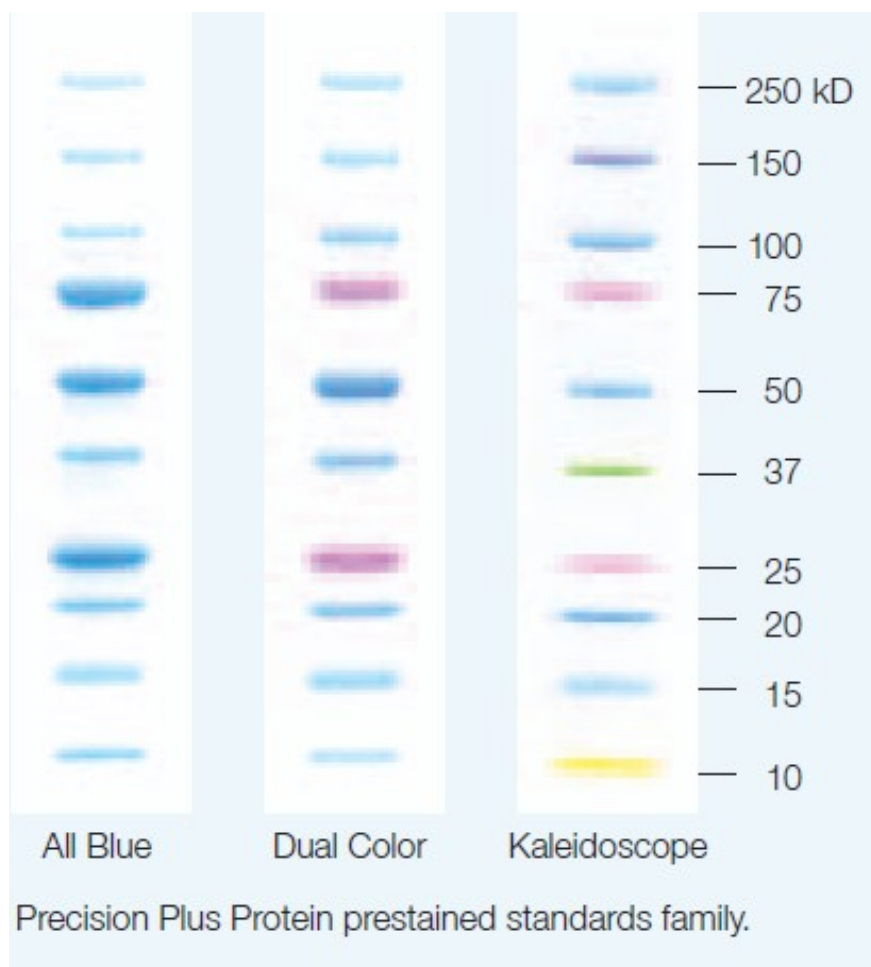
HyperLadder I (Bioline)



**1% agarose gel
5µl per lane**

Protein standards

Precision Plus Protein™ Standards (Bio-Rad)



SigmaMarkers (Both high and low ranges used throughout this work)

Proteins	Molecular Weight	High Range (S8320)	Wide Range (S8445)	Low Range (M3913)
Myosin from rabbit muscle	200,000	X	X	
β -Galactosidase from <i>E. coli</i>	116,000	X	X	
Phosphorylase b from rabbit muscle	97,000	X	X	
Albumin, bovine serum	66,000	X	X	X
Glutamic Dehydrogenase from bovine liver	55,000	X	X	
Ovalbumin from chicken egg	45,000	X	X	X
Glyceraldehyde-3-phosphate Dehydrogenase from rabbit muscle	36,000	X	X	X
Carbonic Anhydrase from bovine erythrocytes	29,000		X	X
Trypsinogen from bovine pancreas	24,000		X	X
Trypsin Inhibitor from soybean	20,000		X	X
α -Lactalbumin from bovine milk	14,200		X	X
Aprotinin from bovine lung	6,500		X	X